TRANSACTIONS

OF THE

AMERICAN INSTITUTE

OF

ELECTRICAL ENGINEERS

VOLUME 74

1955

PART I: COMMUNICATION AND ELECTRONICS

PUBLISHED BY THE
AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
33 WEST 39TH STREET
NEW YORK 18, N. Y., U. S. A.

Copyright 1955 by the American Institute of Electrical Engineers

Printed in U. S. A. by The Mack Printing Company, Easton, Pa.

Preface

The AIEE *Transactions* for 1955 (volume 74) is published in three parts: Part I. Communication and Electronics; Part II. Applications and Industry; and Part III. Power Apparatus and Systems. The papers in each of the three parts are classified according to subject matter as follows:

Part I. Communication and Electronics

- Communication Switching Systems
- Communication Theory
- Radio Communications Systems
- Special Communications Applications
- Telegraph Systems
- Television and Aural Broadcasting Systems
- Wire Communications Systems
- Basic Sciences
- Computing Devices
- Dielectrics
- Electrical Techniques in Medicine and Biology
- Electronics
- Instruments and Measurements
- Magnetic Amplifiers
- Metallic Rectifiers
- Nucleonics
- Solid State Devices

Part II. Applications and Industry

- Air Transportation
- Domestic and Commercial Applications
- Land Transportation
- Marine Transportation
- Production and Application of Light
- Chemical Industry
- Electric Heating
- Electric Welding
- Feedback Control Systems
- General Industry Applications
- Industrial Control
- Industrial Power Rectifiers
- Industrial Power Systems
- Mining and Metal Industry
- Petroleum Industry

Part III. Power Apparatus and Systems

- Carrier Current
- Insulated Conductors
- Power Generation
- Protective Devices
- Relays
- Rotating Machinery
- Substations
- Switchgear
- System Engineering
- Transformers
- Transmission and Distribution

- Education
- Safety
- Research

Each part has been indexed separately in the back of that particular part. The three parts are not cross-referenced; hence the user should determine first whether the subject matter of the paper desired is in the field of communication and electronics, applications and industry, or power apparatus and systems. Papers are listed in the subject index under several key words in the titles. The original numbers assigned to the papers are given in the author index. Volume 74 contains the technical papers and related discussions presented at these meetings:

1. Winter General Meeting, New York, N. Y., January 31–February 4, 1955.
2. Southern District Meeting, St. Petersburg, Fla., April 13–15, 1955.
3. Middle Eastern District Meeting, Columbus Ohio, May 4–6, 1955.
4. Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955.
5. Pacific General Meeting, Butte, Mont., August 15–17, 1955.
6. Fall General Meeting, Chicago, Ill., October 3–7, 1955.

Statements and opinions given in the papers and discussions published in *Transactions* are the expressions of the contributors for which the American Institute of Electrical Engineers assumes no responsibility.

Vol.
Pt. 1
1955

Communication and Electronics Index for 1955

1. Technical Subject Index

A-C Electromotive-Force Transducer and Its Application to Null-Type Recorders, An Angular Position. Guptill, McCarty....610–13
A-C Operated Repeater for 12-Channel Open-Wire Carrier Systems, A Pole-Mounted. Searle, Schaefer....791–5; disc. 795
A-C Servo Applications, Transistor Demodulator for High Performance Magnetic Amplifiers in. Decker....121–3
A-C Windings, A Circuit for Measuring the Resistance of Energized. Seely....214–17; disc. 217
Accurate Electronic Tracer for Dynamic Characteristics of Magnetic Materials, An. Kittl....407–18
(ACSR) for Railroad Signal and Communications Circuits, Aluminum Conductor Steel-Reinforced. Adams....627–9
Admittance, Rapid Measurement of Impedance and. Salzberg, Marini....180–6
Aging Characteristics as Measured by Shearing Modulus, Insulation. Whitman, Scheideler.232–8; disc. 238
Aging Tests on Flexible Sheet Insulation, Significant Factors in Thermal. Dakin, Philofsky, Divens....289–92; disc. 292
Air-borne Temperature Indicator, An. Clark, Amey, Mergner....1–5
Air Traffic Control, Data Transfer and Display Equipment for a Proposed System of. Fenimore..145–9
Air Traffic Control Information, An Electronic System for. Kalb....152–5
All-Magnetic Audio Amplifier System, An. Suozzi, Hooper....297–300; disc. 300
Aluminum Conductor Steel-Reinforced (ACSR) for Railroad Signal and Communications Circuits. Adams....627–9
Aluminum Die Castings for Carrier Telephone Systems. Pedersen....434–9
Amplifier, A Fast Response Magnetic Servo. Kallander....49–54
Amplifier System, An All-Magnetic Audio. Suozzi, Hooper....297–300; disc. 300
Amplifiers, On the Equivalent Circuits of Linear. Vallese....691–4
Amplitude-Modulated Signals, Transfer Functions for. Hellerman....729–31
An Elliptical Polarization Synthesizer. Friedman....339–43; disc. 343
Analogue Divider, A Static Magnetic Transistor. Schaefer....652–5; disc. 655
Analogues Through Conformal Mapping, Improving Field. Farr, Keen....395–400; disc. 400
Analysis and Design of a Magnetic Frequency Multiplier. Smith, Salihi....99–106
Analysis of a Triode Oscillator with Losses Associated with the Inductive Branch and Signal Applied in Grid Circuit. Mostafa....777–81
Analysis of Nonlinear Coupled Circuits. Part II. Ku....439–43
Analysis of Optimum Core Configurations for Magnetic Amplifiers Using a Simplified Approach, An. Bedford, Willis, Dodson....62–70
Analyzer Operation, Computer for Automatizing Network. Kusko, Heller....208–13; disc. 213
Angular Position A-C Electromotive-Force Transducer and Its Application to Null-Type Recorders, An. Guptill, McCarty....610–13
Application of Perpendicularly Superposed Magnetic Fields, The. Heartz, Buelteman..655–9; disc. 659
Appraisal of Delays in Gate-Type Operation, The. Molnar....475–85
Asymmetrical Type, Magnetic Amplifier Sequential Circuits of the. Fitzgerald....685–91
Asymmetry Zero-Drift Errors in Magnetic Servo Amplifiers, Elimination of. Geyger....239–44; disc. 244
(Atlantic) A Transatlantic Telephone Cable. Kelly, Radley, Gilman, Halsey....124–36; disc. 136
Atmospheric Electrostatic Potential, Measuring and Recording. Carroll, Hammond, Stewart..517–20
Audio Amplifier System, An All-Magnetic. Suozzi, Hooper....297–300; disc. 300
Audio-Frequency Range, An Envelope Delay-Measuring Instrument in the. Cannon....710–17
Audio-Frequency Transistor Power Supply, Operational Magnetic Amplifier with. Decker, Gourash.490–6
Audio Telegraph Carrier Terminal, A New. Higgins, Gilbert....58–62
Automatic Field Plotter, An. Gelfand, Shinn, Tuteur....73–8
Automation by Signal Control, Textile. Casper..176–80
Automatizing Network Analyzer Operation, Computer for. Kusko, Heller....208–13; disc. 213

B

Balance for the Measurement of Small Torques, A Precision Torque. Matthys....485–90
Basic Problems in Film Pickup for Television Broadcasting. Gore....462–6
(Basic Sciences) A Continuous-Flow Moisture Detector. Beavers, Shimanski, Timpane.....361–6; disc. 366
(Basic Sciences) A Method for Evaluation of Thermal Stability of Magnet Wire Enamel. Currin, Dexter....227–31; disc. 231
(Basic Sciences) A Method of Solving Mathieu's Equation. Smith....520–5
(Basic Sciences) A Supplement to the Brune Synthesis. Reza....85–90
(Basic Sciences) An Automatic Field Plotter. Gelfand, Shinn, Tuteur....73–8
(Basic Sciences) Analysis of a Triode Oscillator with Losses Associated with the Inductive Branch and Signal Applied in Grid Circuit. Mostafa...777–81
(Basic Sciences) Analysis of Nonlinear Coupled Circuits. Part II. Ku....439–43
(Basic Sciences) Determination of Thermal Life of Enameled Wire by Laboratory Methods. Sattler....70–3
(Basic Sciences) Dielectric Failure of Volume and Surface Types. Chapman, Frisco, Smith....349–54; disc. 354
(Basic Sciences) Effects of Transverse Compressional Stress on Magnetic Laminations. Fischell...190–3
(Basic Sciences) Electric Breakdown of Perfluorocarbon Vapors and Their Mixture With Nitrogen. Berberich, Works, Lindsay....660–5; disc. 665
(Basic Sciences) Equations for the Inductance and Short-Circuit Forces of 3-Phase Busses Comprised of 120-Degree Angles. Maier, Higgins..334–8; disc. 338
(Basic Sciences) Functional Evaluation of Magnetic Wire Insulation. Brancato, Phillips...239–6; disc. 297
(Basic Sciences) Gas and Moisture Equilibriums in Transformer Oil. Kaufman, Shimanski, MacFadyen....312–18; disc. 318
(Basic Sciences) Heat-Resistant Insulation Systems for Motors. Herman, Mathes....561–5
(Basic Sciences) Heat-Resistant Polyester Magnet Wire. Gilliam, Boldebuck, Elliott....557–60
(Basic Sciences) Improving Field Analogues Through Conformal Mapping. Farr, Keen....395–400; disc. 400
(Basic Sciences) Insulation Aging Characteristics as Measured by Shearing Modulus. Whitman, Scheideler....232–8; disc. 238
(Basic Sciences) Junction Transistors Used as Switches. Bright....111–21
(Basic Sciences) Magnetic Flux Density Measurements in Rotating Machines. Ley, Charmatz.....525–8
(Basic Sciences) Phenomena Accompanying Transient Low-Voltage Discharges in Liquid Dielectrics. Williams, Smith....164–9
(Basic Sciences) Properties of Silicon Power Rectifiers. Lusco....106–11
(Basic Sciences) Recommendations of IEC Technical Committee No. 24 on Electric and Magnetic Magnitudes and Units. Smith....37–9
(Basic Sciences) Relation of D-C Magnetic Properties of Oriented 48-Percent Nickel-Iron to Magnetic Amplifier Performance. Ward, Littmann...422–7
(Basic Sciences) Significant Factors in Thermal Aging Tests on Flexible Sheet Insulation. Dakin, Philofsky, Divens....289–92; disc. 292
(Basic Sciences) Some Electrical and Thermal Characteristics of Helium and Sulfur-Hexafluoride Mixtures. Cobine....318–21; disc. 321
(Basic Sciences) Some Properties of Magnetic Fluids. Coolidge, Halberg....149–52
(Basic Sciences) Surface and Volume Dielectric Losses. Chapman, Blickley, Szymkowiak..343–8; disc. 348
(Basic Sciences) Switching Transistors Used as a Substitute for Mechanical Low-Level Choppers. Kruper....141–4
(Basic Sciences) The Dielectric Behavior of Some Fluorogases and Their Mixtures. Camilli, Liao, Plump....637–42; disc. 642
(Basic Sciences) The Propagation Mechanism of Impulse Creepage Discharges Over Oil-Immersed Surfaces. Anderson, Liao....218–26; disc. 226
(Basic Sciences) The Underwater Spark; A Photographic Light Source of High Intrinsic Brilliance. Early, Martin....788–90
(Basic Sciences) Theory of Magnetic Cross-Valves. Higa....204–07; disc. 207
(Basic Sciences) Transfer Functions for Amplitude-Modulated Signals. Hellerman....729–31
(Basic Sciences) What Is a Minimum Phase Network? Balabanian, LePage....785–7; disc. 787
Behavior of Some Fluorogases and Their Mixtures, The Dielectric. Camilli, Liao, Plump.637–42; disc. 642
Bias and Distortion Measuring Instrument, A Portable Telegraph. Cannon....401–06; disc. 406
Bistable Operation, Magnetic Amplifiers in. Finzi, Feth....592–8
Board for Portable Watthour Meter Standards, A New Test. Levitsky....367–70; disc. 370
Breakdown of Perfluorocarbon Vapors and Their Mixture With Nitrogen, Electric. Berberich, Works, Lindsay....660–5; disc. 665
Broad-Band and Pulse Transformer Cores, Presentation of Data on. Gillette, Oshima, Henderson, Rowe....755–63
Brune Synthesis, A Supplement to the. Reza....85–90
Bushings, Progress in the Evaluation of Solid-Core High-Voltage. Brustle, Johnston, Scheideler....387–91; disc. 391
Busses Comprised of 120-Degree Angles, Equations for the Inductance and Short-Circuit Forces of 3-Phase. Maier, Higgins....334–8; disc. 338

C

Cable, A Transatlantic Telephone. Kelly, Radley, Gilman, Halsey....124–36; disc. 136
Cable Carrier and 12-Channel Open-Wire Carrier Systems, Combined Operation of 24-Channel. Searle, Schaefer....770–6; disc. 777
Cables, Shielding of Communication. Gooding, Slade....378–86; disc. 386
Carrier and 12-Channel Open-Wire Carrier Systems, Combined Operation of 24-Channel Cable. Searle, Schaefer....770–6; disc. 777
Carrier Systems, A Pole-Mounted A-C Operated Repeater for 12-Channel Open-Wire. Searle, Schaefer....791–5; disc. 795
Carrier Systems, Frequency-Shift Signaling Circuit for 45-Type. Appert, Caruthers....613–19
Carrier Systems in South Africa, Open-Wire. Boyce....272–81
Carrier Telephone Systems, Aluminum Die Castings for. Pedersen....434–9
Carrier Terminal, A New Audio Telegraph. Higgins, Gilbert....58–62
Cathode Counting Circuits, Cold-. Foote....161–4
Cathode Followers, On the Optimum Design of. Vallese....5–8
Cavity Filters for Multiplexing of 900-Megacycle Radio Relay Systems, Coaxial. Kebby....261–7
Channel Cable Carrier and 12-Channel Open-Wire Carrier Systems, Combined Operation of 24-. Searle, Schaefer....770–6; disc. 777
Characteristics as Measured by Shearing Modulus, Insulation Aging. Whitman, Scheideler....232–8; disc. 238
Characteristics of Helium and Sulfur-Hexafluoride Mixtures, Some Electrical and Thermal. Cobine....318–21; disc. 321
Characteristics of Magnetic Materials, An Accurate Electronic Tracer for Dynamic. Kittl....407–18
Choppers, Switching Transistors Used as a Substitute for Mechanical Low-Level. Kruper....141–4
Circuit, A Multistable Transistor. Henle....568–71
Circuit Design, Notes on the Use of Screen-to-Plate Transconductance in Multigrid Tube. Pullen....648–51; disc. 651

ol. 74
t. 1
955

Circuit Drawing, An Improved Detached Contact Type of Schematic. Meyer........505–13
Circuit Equations, Formulation of Magnetic Amplifier. Schaefer........619–27
Circuit for Electronic Power Factor Meters and Wattmeters, and a Novel Division Device, A Novel. Ahmed........194–203; disc. 203
Circuit for 45-Type Carrier Systems, Frequency-Shift Signaling. Appert, Caruthers........613–19
Circuit for Measuring the Resistance of Energized A-C Windings, A. Seely........214–17; disc. 217
Circuits, A New Telegraph Serviceboard Using Electronic. Davey, Hanley, Purvis........30–7
Circuits, Aluminum Conductor Steel-Reinforced (ACSR) for Railroad Signal and Communications. Adams........627–9
Circuits, Analysis of Nonlinear Coupled. Part II. Ku........439–43
Circuits, Cold-Cathode Counting. Foote........161–4
Circuits, Engineering Multistage Diode Logic. Yokelson, Ulrich........466–75
Circuits, Equipment for Interisland Radio. Walker, Nolan, Hoover........576–83
Circuits, Minimizing and Mapping Sequential. Bennett........443–7; disc. 796
Circuits of Linear Amplifiers, On the Equivalent. Vallese........691–4
Circuits of the Asymmetrical Type, Magnetic Amplifier Sequential. Fitzgerald........685–91
Closed-Circuit Television, Developments in. Kraus........704–09
Coaxial Cavity Filters for Multiplexing of 900-Megacycle Radio Relay Systems. Kebby........261–7
Cold-Cathode Counting Circuits. Foote........161–4
Color Printing, Electronic Computer for. Rose..268–72
Combined Operation of 24-Channel Cable Carrier and 12-Channel Open-Wire Carrier Systems. Searle, Schaefer........770–6; disc. 777
Commercial Oil, Empirical Equations for the 60-Cycle Spark-over Gradients in. Gauster........301–07; disc. 307
Committee No. 24 on Electric and Magnetic Magnitudes and Units, Recommendations of IEC Technical. Smith........37–9
(Committee Report) Standards for the Presentation of Data on Magnetic Amplifier Core Materials. (55-540)........598–9
(Communication) A Pole-Mounted A-C Operated Repeater for 12-Channel Open-Wire Carrier Systems. Searle, Schaefer........791–5; disc. 795
(Communication) A Transatlantic Telephone Cable. Kelly, Radley, Gilman, Halsey..124–36; disc. 136
(Communication) An Improved Detached Contact Type of Schematic Circuit Drawing. Meyer..505–13
Communication Cables, Shielding of. Gooding, Slade........378–86; disc. 386
(Communication) Development of a Wire Contact Relay. Markle........742–5
(Communication) Engineering Multistage Diode Logic Circuits. Yokelson, Ulrich........466–75
(Communication) Minimizing and Mapping Sequential Circuits. Bennett........443–7; disc. 796
(Communication) Roll-Welding Precious Metals for Telephone Contacts. Quinlan........763–6
(Communication) The Appraisal of Delays in Gate-Type Operation. Molnar........475–85
(Communication) Very High Frequency Radio Link Between Puerto Rico and the Virgin Islands. McSweeny........781–5
Communications Circuits, Aluminum Conductor Steel-Reinforced (ACSR) for Railroad Signal and. Adams........627–9
Communications System for the New York State Thruway, A. DeWire, Evans........253–61
Compressional Stress on Magnetic Laminations, Effects of Transverse. Fischell........190–3
Computer for Automatizing Network Analyzer Operation. Kusko, Heller........208–13; disc. 213
Computer for Color Printing, Electronic. Rose..268–72
(Computing Devices) A Multistable Transistor Circuit. Henle........568–71
(Computing Devices) A Sine Function Resistor. Nielsen, Roland........447–50
(Computing Devices) A Static Magnetic Transistor Analogue Divider. Schaefer. (55-451)........652–5; disc. 655
(Computing Devices) An Electronic System for Air Traffic Control Information. Kalb........152–5
(Computing Devices) Data Transfer and Display Equipment for a Proposed System of Air Traffic Control. Fenimore........145–9
(Computing Devices) Four-Quadrant Multiplication with Transistors and Magnetic Cores. Van Allen........643–7; disc. 647
(Computing Devices) The Design of the IBM Type 702 System. Bashe, Jackson, Mussell, Winger..695–704
Conductor Steel-Reinforced (ACSR) for Railroad Signal and Communications Circuits, Aluminum. Adams........627–9
Configurations for Magnetic Amplifiers Using a Simplified Approach, An Analysis of Optimum Core. Bedford, Willis, Dodson........62–70
Conformal Mapping, Improving Field Analogues Through. Farr, Keen........395–400; disc. 400
Contact Relay, Development of a Wire. Markle..742–5
Contact Type of Schematic Circuit Drawing, An Improved Detached. Meyer........505–13
Continuous-Flow Moisture Detector, A. Beavers, Shimanski, Timpane........361–6; disc. 366
Control of D-C Motors, Magnetic Amplifier. Kusko, Nelson........326–33, disc. 333
Control, Textile Automation by Signal. Casper.176–80
Converter Having an Output Frequency Proportional to the D-C Input Voltage, A Switching Transistor D-C to A-C. Royer........322–4; disc. 324
Converters for Polyphase Power Supplies, Phase Locking of Switching-Transistor. Milnes........587–92
Core Configurations for Magnetic Amplifiers Using a Simplified Approoch, An Analysis of Optimum. Bedford, Willis, Dodson........62–70
Core Materials, Standards for the Presentation of Data on Magnetic Amplifier. (Committee Report).598–9
Core Transformer as a Pickup for Industrial Telemetering Systems, The Movable. Hornfeck, Wermelskirchen........454–62
Cores, Presentation of Data on Broad-Band and Pulse Transformer. Gillette, Oshima, Henderson, Rowe........755–63
Counters in X-Ray Diffraction, Use of Proportional. Laird, Zunick........734–7
Counting Circuits, Cold-Cathode. Foote........161–4
Coupled Circuits, Analysis of Nonlinear. Part II. Ku........439–43
Creepage Discharges Over Oil-Immersed Surfaces, The Propagation Mechanism of Impulse. Anderson, Liao........218–26; disc 226
Cross-Valves, Theory of Magnetic. Higa........204–07; disc. 207
Currents of Double-Way and Double-Wye Rectifiers, Regulator Curves and Transient. Jensen, Rettig........281–7; disc. 287
Curves and Transient Currents of Double-Way and Double-Wye Rectifiers, Regulator. Jensen, Rettig........281–7; disc. 287

D

D-C Dielectric Testing in the Field, The Recognition of Possible Measurement Errors in. Ross, Curdts........630–5; disc. 635
D-C Input Voltage, A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the. Royer........322–4; disc. 324
D-C Magnetic Properties of Oriented 48-Percent Nickel-Iron to Magnetic Amplifier Performance, Relation of. Ward, Littmann........422–7
D-C Motors, Magnetic Amplifier Control of. Kusko, Nelson........326–33; disc. 333
Data from a Poisson Distribution, Machine Testing for Deviation of. Hadden........155–7
Data on Broad-Band and Pulse Transformer Cores, Presentation of. Gillette, Oshima, Henderson, Rowe........755–63
Data on Magnetic Amplifier Core Materials, Standards for the Presentation of. (Committee Report).598–9
Data Transfer and Display Equipment for a Proposed System of Air Traffic Control. Fenimore...145–9
Decicycle Magnetic Amplifier Systems for Servo Applications. Johnson, Rauch........667–72; disc. 672
Delay Device Utilizing Transistors, A High-Accuracy Static Time. Pittman........54–8
Delay-Measuring Instrument in the Audio-Frequency Range, An Envelope. Cannon........710–17
Delays in Gate-Type Operation, The Appraisal of. Molnar........475–85
Demodulator for High-Performance Magnetic Amplifiers in A-C Servo Applications, Transistor. Decker........121–3
Density Measurements in Rotating Machines, Magnetic Flux. Ley, Charmatz........525–8
Design of a Magnetic Frequency Multiplier, Analysis and. Smith, Salihi........99–106
Design of a New 2-Element Single Disk Polyphase Meter, The. Schmidt........513–16
Design of Cathode Followers, On the Optimum. Vallese........5–8
Design of Saturating Reactor Magnetic Pulsers, Economic. Mathias, Williams........169–71
Design of the IBM Type 702 System, The. Bashe, Jackson, Mussell, Winger........695–704
Detached Contact Type of Schematic Circuit Drawing, An Improved. Meyer........505–13
Detector, A Continuous-Flow Moisture. Beavers, Shimanski, Timpane........361–6; disc. 366
Determination of Thermal Life of Enameled Wire by Laboratory Methods. Sattler........70–3
Development of a Wire Contact Relay. Markle...742–5
Developments in Closed-Circuit Television. Kraus........704–09
Device Utilizing Transistors, A High-Accuracy Static Time Delay. Pittman........54–8
Die Castings for Carrier Telephone Systems, Aluminum. Pedersen........434–9
Dielectric Amplifier, Voltage Gain of a Resonant. Sack, Penney........428–34
Dielectric Behavior of Some Fluorogases and Their Mixtures, The. Camilli, Liao, Plump........637–42; disc. 642
Dielectric Failure of Volume and Surface Types. Chapman, Frisco, Smith........349–54; disc. 354
Dielectric Losses, Surface and Volume. Chapman, Blickley, Szymkowiak........343–8; disc. 348
Dielectric Solids, The Effects of High-Energy Gamma Radiation on. Klein, Mannal........723–9
Dielectric Testing in the Field, The Recognition of Possible Measurement Errors in D-C. Ross, Curdts........630–5; disc. 635
Dielectric Tests on a Large Hydrogenerator, Report of. Duke, Ross, Johnson........673–8; disc. 678
Dielectrics, Phenomena Accompanying Transient Low-Voltage Discharges in Liquid. Williams, Smith........164–9
(Dielectrics) The Effect of Reactor Irradiation on Electrical Insulation. Pigg, Bopp, Sisman, Robinson........717–23
Diffraction, Use of Proportional Counters in X-Ray. Laird, Zunick........734–7
Diode Logic Circuits, Engineering Multistage. Yokelson, Ulrich........466–75
Direct-Acting Recording Instruments, A New Line of. Lunas, Nycz........679–84; disc. 684
Discharges in Liquid Dielectrics, Phenomena Accompanying Transient Low-Voltage. Williams, Smith........164–9
Discharges Over Oil-Immersed Surfaces, The Propagation Mechanism of Impulse Creepage. Anderson, Liao........218–26; disc. 226
Disk Type of Tachometer Indicator, Temperature Errors in a Dragmagnet Eddy-Current. Akeley, Fraizer........418–22
Distortion, A Symmetrical Transistor Oscillator with Low Second-Harmonic. Grim...767–70; disc. 770
Distortion Measuring Instrument, A Portable Telegraph Bias and. Cannon........401–06; disc. 406
Distribution, Machine Testing for Deviation of Data from a Poisson. Hadden........155–7
Divider, A Static Magnetic Transistor Analogue. Schaefer........652–5; disc. 655
Division Device, A Novel Circuit for Electronic Power Factor Meters and Wattmeters, and a Novel. Ahmed........194–203; disc. 203
Double-Way and Double-Wye Rectifiers, Regulator Curves and Transient Currents of. Jensen, Rettig........281–7; disc. 287
Dragmagnet Eddy-Current Disk Type of Tachometer Indicator, Temperature Errors in a. Akeley, Fraizer........418–22
Drawing, An Improved Detached Contact Type of Schematic Circuit. Meyer........505–13
Dynamic Characteristics of Magnetic Materials, An Accurate Electronic Tracer for. Kittl......407–18
Dynamic Magnetic Properties of Toroidal Cores, Influence of ID/OD Ratio on Static and. Roberts, Van Nice........599–607

E

Economic Design of Saturating Reactor Magnetic Pulsers. Mathias, Williams........169–71
Eddy-Current Disk Type of Tachometer Indicator, Temperature Errors in a Dragmagnet. Akeley, Fraizer........418–22
Eddy-Current Mutual-Inductance Transducers With High-Conductivity Reference Plates. Joseph, Newman........39–49
Effect of Reactor Irradiation on Electrical Insulation, The. Pigg, Bopp, Sisman, Robinson........717–23
Effects of High-Energy Gamma Radiation on Dielectric Solids. Klein, Mannal........723–9
Effects of Transverse Compressional Stress on Magnetic Laminations. Fischell........190–3
Electric and Magnetic Magnitudes and Units, Recommendations of IEC Technical Committee No. 24 on. Smith........37–9
Electric Breakdown of Perfluorocarbon Vapors and Their Mixture With Nitrogen. Berberich, Works, Lindsay........660–5; disc. 665
Electric System, Operation of a Nuclear Power Plant on an Integrated. Wilson........751–4; disc. 754
Electrical and Thermal Characteristics of Helium and Sulfur-Hexafluoride Mixtures, Some. Cobine........318–21; disc. 321
Electrical Insulation, The Effect of Reactor Irradiation on. Pigg, Bopp, Sisman, Robinson........717–23
Electromotive-Force Transducer and Its Application to Null-Type Recorders, An Angular Position A-C. Guptill, McCarty........610–13
Electronic Circuits, A New Telegraph Serviceboard Using. Davey, Hanley, Purvis........30–7
Electronic Computer for Color Printing. Rose...268–72

Electronic Power Factor Meters and Wattmeters, and a Novel Division Device, A Novel Circuit for. Ahmed ....194–203; disc. 203
Electronic System for Air Traffic Control Information, An. Kalb............152–5
Electronic Tracer for Dynamic Characteristics of Magnetic Materials, An Accurate. Kitil.......407–18
(Electronics) A Symmetrical Transistor Oscillator with Low Second-Harmonic Distortion. Grim............767–70; disc. 770
(Electronics) Cold-Cathode Counting Circuits. Foote 161–4
(Electronics) Machine Testing for Deviation of Data from a Poisson Distribution. Hadden.......155–7
(Electronics) Notes on the Use of Screen-to-Plate Transconductance in Multigrid Tube Circuit Design. Pullen............648–51; disc. 651
(Electronics) On the Equivalent Circuits of Linear Amplifiers. Vallese............691–4
(Electronics) On the Optimum Design of Cathode Followers. Vallese............5–8
(Electronics) Science, Statistics, and Reliability Engineering. Bear............82–5
(Electronics) Shielding of Communication Cables. Gooding, Slade............378–86; disc. 386
Electronics, Ticketfax. Jelinek............565–8.
(Electronics) Use of Proportional Counters in X-Ray Diffraction. Laird, Zunick............734–7.
Electrostatic Potential, Measuring and Recording Atmospheric. Carroll, Hammond, Stewart...517–20
Elimination of Asymmetry Zero-Drift Errors in Magnetic Servo Amplifiers. Geyger.....239–44; disc. 244
Elliptical Polarization Synthesizer, An. Friedman............339–43; disc. 343
Empirical Equations for the 60-Cycle Spark-over Gradients in Commercial Oil. Gauster..301–07; disc. 307
Enamel, A Method for Evaluation of Thermal Stability of Magnet Wire. Currin, Dexter..227–31; disc. 231
Enameled Wire by Laboratory Methods, Determination of Thermal Life of. Sattler............70–3
Energized A-C Windings, A Circuit for Measuring the Resistance of. Seely............214–17; disc. 217
Engineering Multistage Diode Logic Circuits. Yokelson, Ulrich............466–75
Engineering Problems and Performance Record of the Hawaiian Interisland Radio Network. Farinon, Walker............571–6
Engineering, Science, Statistics, and Reliability. Bear 82–5
Envelope Delay-Measuring Instrument in the Audio-Frequency Range, An. Cannon............710–17
Equation, A Method of Solving Mathieu's. Smith 520–5
Equations for the Inductance and Short-Circuit Forces of 3-Phase Busses Comprised of 120-Degree Angles. Maier, Higgins............334–8; disc. 338
Equations for the 60-Cycle Spark-over Gradients in Commercial Oil, Empirical. Gauster............301–07; disc. 307
Equations, Formulation of Magnetic Amplifier Circuit. Schaefer............619–27
Equilibriums in Transformer Oil, Gas and Moisture. Kaufman, Shimanski, MacFadyen............312–18; disc. 318
Equipment for a Proposed System of Air Traffic Control, Data Transfer and Display. Fenimore......145–9
Equipment for Interisland Radio Circuits. Walker, Nolan, Hoover............576–83
Equivalent Circuits of Linear Amplifiers, On the. Vallese............691–4
Errors in a Dragmagnet Eddy-Current Disk Type of Tachometer Indicator, Temperature. Akeley, Fraizer............418–22
Errors in D-C Dielectric Testing in the Field, The Recognition of Possible Measurement. Ross, Curdts............630–5; disc. 635
Errors in Magnetic Servo Amplifiers, Elimination of Asymmetry Zero-Drift. Geyger...239–44; disc. 244
Evaluation of Magnetic Amplifier Response, Oscillographic Techniques for the. Critchlow...607–10
Evaluation of Magnetic Wire Insulation, Functional. Brancato, Phillips............293–6; disc. 297
Evaluation of Solid-Core High-Voltage Bushings, Progress in the. Brustle, Johnston, Scheideler............387–91; disc. 391
Evaluation of Thermal Stability of Magnet Wire Enamel, A Method for. Currin, Dexter..227–31; disc. 231

F

Failure of Volume and Surface Types, Dielectric. Chapman, Frisco, Smith............349–54; disc. 354
Fast Response Magnetic Servo Amplifier, A. Kallander 49–54
Field Analogues Through Conformal Mapping, Improving. Farr, Keen............395–400; disc. 400
Field Plotter, An Automatic. Gelfand, Shinn, Tuteur 73–8
Field Tests on a 100-Megawatt Rectifier Installation. Dillard, Kiefer, Hague............245–52; disc. 252
Film Pickup for Television Broadcasting, Basic Problems in. Gore............462–6
Filters for Multiplexing of 900-Megacycle Radio Relay Systems, Coaxial Cavity. Kebby............261–7
Flexible Sheet Insulation, Significant Factors in Thermal Aging Tests on. Dakin, Philofsky, Divens............289–92; disc. 292
Flow Moisture Detector, A Continuous-. Beavers, Shimanski, Timpane............361–6; disc. 366
Fluids, Some Properties of Magnetic. Coolidge, Halberg 149–52
Fluorogases and Their Mixtures, The Dielectric Behavior of Some. Camilli, Liao, Plump............637–42; disc. 642
Flux Density Measurements in Rotating Machines, Magnetic. Ley, Charmatz............525–8
Force Transducer and Its Application to Null-Type Recorders, An Angular Position A-C Electromotive-. Guptill, McCarty............610–13
Forces of 3-Phase Busses Comprised of 120-Degree Angles, Equations for the Inductance and Short-Circuit. Maier, Higgins.........334–8; disc. 338
Formulation of Magnetic Amplifier Circuit Equations. Schaefer............619–27
Four-Quadrant Multiplication with Transistors and Magnetic Cores. Van Allen......643–7; disc. 647
Frequency Magnetic-Coupled Multivibrator, A Variable. Van Allen............356–61; disc. 361
Frequency Meter, A Transducer-Type. Smith..497–500
FM Receivers, Impulse Noise in Narrow-Band. Lapin, Suran............450–4
Frequency Multiplier, Analysis and Design of a Magnetic. Smith, Salihi............99–106
Frequency Proportional to the D-C Input Voltage, A Switching Transistor D-C to A-C Converter Having an Output. Royer............322–4; disc. 324
Frequency Range, An Envelope Delay-Measuring Instrument in the Audio-. Cannon........710–17
Frequency-Shift Signaling Circuit for 45-Type Carrier Systems. Appert, Caruthers............613–19
Frequency Transistor Power Supply, Operational Magnetic Amplifier with Audio-. Decker, Gourash............490–6
Full-Wave Magnetic Amplifier Output Stage, A New. Barnhart............139–41
Full Wave Reversible Polarity Half Cycle Response Magnetic Amplifiers. House............541–52
Fully Selective Telemetering System Employing Telegraph Facilities, A. Smith, Forrest............373–7; disc. 377
Function Resistor, A Sine. Nielsen, Roland.....447–50
Functional Evaluation of Magnetic Wire Insulation. Brancato, Phillips............293–6; disc. 297
Functions for Amplitude-Modulated Signals, Transfer. Hellerman............729–31

G

Gain of a Resonant Dielectric Amplifier, Voltage. Sack, Penney............428–34
Gamma Radiation on Dielectric Solids, The Effects of High-Energy. Klein, Mannal............723–9
Gas and Moisture Equilibriums in Transformer Oil. Kaufman, Shimanski, MacFadyen............312–18; disc. 318
Gate-Type Operation, The Appraisal of Delays in. Molnar............475–85
Gradients in Commercial Oil, Empirical Equations for the 60-Cycle Spark-over. Gauster............301–07; disc. 307
Grid Circuit, Analysis of a Triode Oscillator with Losses Associated with the Inductive Branch and Signal Applied in. Mostafa............777–81

H

Half Cycle Response Magnetic Amplifiers, Full Wave Reversible Polarity. House............541–52
Harmonic Distortion, A Symmetrical Transistor Oscillator with Low Second-. Grim..767–70; disc. 770
(Hawaiian) Equipment for Interisland Radio Circuits. Walker, Nolan, Hoover............576–83
Hawaiian Interisland Radio Network, Engineering Problems and Performance Record of the. Farinon, Walker............571–6
Hawaiian Telephone Company, The Interisland Radio Network of the. Farinon............529–36
Heat-Resistant Insulation Systems for Motors. Herman, Mathes............561–5
Heat Resistant Polyester Magnet Wire. Gilliam, Boldebuck, Elliott............557–60
Helium and Sulfur-Hexafluoride Mixtures, Some Electrical and Thermal Characteristics of. Cobine ....318–21; disc. 321
Hexafluoride Mixtures, Some Electrical and Thermal Characteristics of Helium and Sulfur-. Cobine ....318–21; disc. 321
High-Accuracy Static Time Delay Device Utilizing Transistors, A. Pittman............54–8
High-Energy Gamma Radiation on Dielectric Solids, The Effects of. Klein, Mannal............723–9
High Frequency Operation of Self-Saturating Magnetic Amplifiers. Collins............500–05
High Frequency Radio Link Between Puerto Rico and the Virgin Islands, Very. McSweeny........781–5
High-Speed Magnetic Amplifier Design, 160,000-Ampere. Rosenstein............90–9
High-Voltage Bushings, Progress in the Evaluation of Solid-Core. Brustle, Johnston, Scheideler............387–91; disc. 391
Hydrogenerator, Report of Dielectric Tests on a Large. Duke, Ross, Johnson............673–8; disc. 678

I

IBM Type 702 System, The Design of the. Bashe, Jackson, Mussell, Winger............695–704
IEC Technical Committee No. 24 on Electric and Magnetic Magnitudes and Units, Recommendations of. Smith............37–9
Impedance and Admittance, Rapid Measurement of. Salzberg, Marini............180–6
Improved Detached Contact Type of Schematic Circuit Drawing, An. Meyer............505–13
Improving Field Analogues Through Conformal Mapping. Farr, Keen............395–400; disc. 400
Impulse Creepage Discharges Over Oil-Immersed Surfaces, The Propagation Mechanism of. Anderson, Liao............218–26; disc. 226
Impulse Noise in Narrow-Band FM Receivers. Lapin, Suran............450–4
Indicator, An Air-borne Temperature. Clark, Amey, Mergner............1–5
Inductance and Short-Circuit Forces of 3-Phase Busses Comprised of 120-Degree Angles, Equations for the. Maier, Higgins............334–8; disc. 338
Inductive Load Instability in Magnetic Amplifiers. Leon, Rosenstein............8–20
Industrial Telemetering Systems, The Movable Core Transformer as a Pickup for. Hornfeck, Wermelskirchen............454–62
Influence of ID/OD Ratio on Static and Dynamic Magnetic Properties of Toroidal Cores. Roberts, Van Nice............599–607
Information, An Electronic System for Air Traffic Control. Kalb............152–5
Input Voltage, A Switching Transistor D–C to A–C Converter Having an Output Frequency Proportional to the D-C. Royer........322–4; disc. 324
Instability in Magnetic Amplifiers, Inductive Load. Leon, Rosenstein............8–20
Instrument, A Portable Telegraph Bias and Distortion Measuring. Cannon............401–06; disc. 406
Instrument in the Audio-Frequency Range, An Envelope Delay-Measuring. Cannon............710–17
Instruments, A New Line of Direct-Acting Recording. Lunas, Nycz............679–84; disc. 684
(Instruments) A New Test Board for Portable Watthour Meter Standards. Levitsky......367–70; disc. 370
(Instruments) A Novel Circuit for Electronic Power Factor Meters and Wattmeters, and a Novel Division Device. Ahmed.........194–203; disc. 203
(Instruments) A Transducer-Type Frequency Meter. Smith............497–500
(Instruments) An Air-borne Temperature Indicator. Clark, Amey, Mergner............1–5
(Instruments) An Elliptical Polarization Synthesizer. Friedman............339–43; disc. 343
(Instruments) Progress in the Evaluation of Solid-Core High-Voltage Bushings. Brustle, Johnston, Scheideler............387–91; disc. 391
(Instruments) Report of Dielectric Tests on a Large Hydrogenerator. Duke, Ross Johnson............673–8; disc. 678
(Instruments) Temperature Errors in a Dragmagnet Eddy-Current Disk Type of Tachometer Indicator. Akeley, Fraizer............418–22
(Instruments) The Design of a New 2-Element Single-Disk Polyphase Meter. Schmidt............513–16
(Instruments) The Movable Core Transformer as a Pickup for Industrial Telemetering Systems. Hornfeck, Wermelskirchen............454–62
(Instruments) The Ring Modulator as a Polarized Rectifier. Hermont............21–30
Insulation Aging Characteristics as Measured by Shearing Modulus. Whitman, Scheideler............232–8; disc. 238
Insulation, Functional Evaluation of Magnetic Wire. Brancato, Phillips............293–6; disc. 297
Insulation, Significant Factors in Thermal Aging Tests on Flexible Sheet. Dakin, Philofsky, Divens............289–92; disc. 292
Insulation Systems for Motors, Heat-Resistant. Herman, Mathes............561–5
Insulation, The Effect of Reactor Irradiation on Electrical. Pigg, Bopp, Sisman, Robinson......717–23
Integrated Electric System, Operation of a Nuclear Power Plant on an. Wilson.......751–4; disc. 754

ol. 74
. 1
955

Interference Measurements, A Method of Making Screen Room. Mortenson, Truax.......746–50; disc. 750
Interisland Radio Circuits, Equipment for. Walker, Nolan, Hoover.......576–83
Inter-Island Radio Network of the Hawaiian Telephone Company, The. Farinon.......529–36
Iron to Magnetic Amplifier Performance, Relation of D–C Magnetic Properties of Oriented 48-Percent Nickel-. Ward, Littmann.......422–7
Irradiation on Electrical Insulation, The Effect of Reactor. Pigg, Bopp, Sisman, Robinson.......717–23

J

Junction Transistors Used as Switches. Bright...111–21

L

Laboratory Methods, Determination of Thermal Life of Enameled Wire by. Sattler.......70–3
Laminations, Effects of Transverse Compressional Stress on Magnetic. Fischell.......190–3
Life of Enameled Wire by Laboratory Methods, Determination of Thermal. Sattler.......70–3
Light Source of High Intrinsic Brilliance, The Underwater Spark; A Photographic. Early, Martin.......788–90
Linear Amplifiers, On the Equivalent Circuits of. Vallese.......691–4
Liquid Dielectrics, Phenomena Accompanying Transient Low-Voltage Discharges in. Williams, Smith.......164–9
Load Instability in Magnetic Amplifiers, Inductive. Leon, Rosenstein.......8–20
Locking of Switching-Transistor Converters for Polyphase Power Supplies, Phase. Milnes.......587–92
Logic Circuits, Engineering Multistage Diode. Yokelson, Ulrich.......466–75
Losses Associated with the Inductive Branch and Signal Applied in Grid Circuit, Analysis of a Triode Oscillator with. Mostafa.......777–81
Losses, Surface and Volume Dielectric. Chapman, Blickley, Szymkowiak.......343–8; disc. 348
Low-Level Choppers, Switching Transistors Used as a Substitute for Mechanical. Kruper.......141–4
Low-Voltage Discharges in Liquid Dielectrics, Phenomena Accompanying Transient. Williams, Smith.......164–9

M

Machine Testing for Deviation of Data from a Poisson Distribution. Hadden.......155–7
Magnet Wire Enamel, A Method for Evaluation of Thermal Stability of. Currin, Dexter.......227–31; disc. 231
Magnet Wire, Heat Resistant Polyester. Gilliam, Boldebuck, Elliott.......557–60
Magnetic Amplifier Circuit Equations, Formulation of. Schaefer.......619–27
Magnetic Amplifier Control of D-C Motors. Kusko, Nelson.......326–33; disc. 333
Magnetic Amplifier Core Materials, Standards for the Presentation of Data on. (Committee Report).......598–9
Magnetic Amplifier Design, 160,000-Ampere High Speed. Rosenstein.......90–9
Magnetic Amplifier Output Stage, A New Full-Wave. Barnhart.......139–41
Magnetic Amplifier Response, Oscillographic Techniques for the Evaluation of. Critchlow.......607–10
Magnetic Amplifier Sequential Circuits of the Asymmetrical Type. Fitzgerald.......685–91
Magnetic Amplifier Systems for Servo Applications, Decicycle. Johnson, Rauch.......667–72; disc. 672
Magnetic Amplifier with Audio-Frequency Transistor Power Supply, Operational. Decker, Gourash.......490–6
(Magnetic Amplifiers) A High-Accuracy Static Time Delay Device Utilizing Transistors. Pittman..54–8
(Magnetic Amplifiers) A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D–C Input Voltage. Royer.......322–4; disc. 324
Magnetic Amplifiers, Full Wave Reversible Polarity Half Cycle Response. House.......541–52
Magnetic Amplifiers, High Frequency Operation of Self-Saturating. Collins.......500–05
Magnetic Amplifiers in A-C Servo Applications, Transistor Demodulator for High Performance. Decker.......121–3
Magnetic Amplifiers in Bistable Operation. Finzi, Feth.......592–8
Magnetic Amplifiers, Inductive Load Instability in. Leon, Rosenstein.......8–20
(Magnetic Amplifiers) Phase Locking of Switching-Transistor Converters for Polyphase Power Supplies. Milnes.......587–92
Magnetic Amplifiers Using a Simplified Approach, An Analysis of Optimum Core Configurations for. Bedford, Willis, Dodson.......62–70
(Magnetic Amplifiers) Voltage Gain of a Resonant Dielectric Amplifier. Sack, Penney.......428–34
Magnetic Audio Amplifier System, An All-. Suozzi, Hooper.......297–300; disc. 300
Magnetic Cores, Four-Quadrant Multiplication with Transistors and. Van Allen.......643–7; disc. 647
Magnetic-Coupled Multivibrator, A Variable Frequency. Van Allen.......356–61; disc. 361
Magnetic Cross-Valves, Theory of. Higa.......204–07; disc. 207
Magnetic Fields, A Survey of Noncontacting Vibration Pickup Using. Clarke.......186–90
Magnetic Fields, The Application of Perpendicularly Superposed. Heartz, Buelteman..655–9; disc. 659
Magnetic Fluids, Some Properties of. Coolidge, Halberg.......149–52
Magnetic Flux Density Measurements in Rotating Machines. Ley, Charmatz.......525–8
Magnetic Frequency Multiplier, Analysis and Design of a. Smith, Salihi.......99–106
Magnetic Laminations, Effects of Transverse Compressional Stress on. Fischell.......190–3
Magnetic Magnitudes and Units, Recommendations of IEC Technical Committee No. 24 on Electric and. Smith.......37–9
Magnetic Materials, An Accurate Electronic Tracer for Dynamic Characteristics of. Kittl.......407–18
Magnetic Properties of Oriented 48-Percent Nickel-Iron to Magnetic Amplifier Performance, Relation of D-C. Ward, Littmann.......422–7
Magnetic Properties of Toroidal Cores, Influence of ID/OD Ratio on Static and Dynamic. Roberts, Van Nice.......599–607
Magnetic Pulsers, Economic Design of Saturating Reactor. Mathias, Williams.......169–71
Magnetic Servo Amplifier, A Fast Response. Kallander.......49–54
Magnetic Servo Amplifiers, Elimination of Asymmetry Zero-Drift Errors in. Geyger.......239–44; disc. 244
Magnetic Transistor Analogue Divider, A Static. Schaefer.......652–5; disc. 655
Magnetic Wire Insulation, Functional Evaluation of. Brancato, Phillips.......293–6; disc. 297
Magnitudes and Units, Recommendations of IEC Technical Committee No. 24 on Electric and Magnetic. Smith.......37–9
Mapping, Improving Field Analogues Through Conformal. Farr, Keen.......395–400; disc. 400
Mapping Sequential Circuits, Minimizing and. Bennett.......443–7
Materials, Standards for the Presentation of Data on Magnetic Amplifier Core. (Committee Report).......598–9
Mathieu's Equation, A Method of Solving. Smith.......520–5
Measurement Errors in D-C Dielectric Testing in the Field, The Recognition of Possible. Ross, Curdts.......630–5; disc. 635
Measurement of Impedance and Admittance, Rapid. Salzberg, Marini.......180–6
Measurement of Small Torques, A Precision Torque Balance for the. Matthys.......485–90
Measurements, A Method of Making Screen Room Interference. Mortenson, Truax.......746–50; disc. 750
(Measurements) An Angular Position A-C Electromotive-Force Transducer and Its Application to Null-Type Recorders. Guptill, McCarty...610–13
(Measurements) Eddy-Current Mutual-Inductance Tranducers with High-Conductivity Reference Plates. Joseph, Newman.......39–49
Measurements in Rotating Machines, Magnetic Flux Density. Ley, Charmatz.......525–8
Measuring and Recording Atmospheric Electrostatic Potential. Carroll, Hammond, Stewart....517–20
Measuring Instrument, A Portable Telegraph Bias and Distortion. Cannon.......401–06; disc. 406
Measuring Instrument in the Audio-Frequency Range, An Envelope Delay-. Cannon.......710–17
Measuring the Resistance of Energized A-C Windings, A Circuit for. Seely.......214–17; disc. 217
Mechanical Low-Level Choppers, Switching Transistors Used as a Substitute for. Kruper.......141–4
Mechanism of Impulse Creepage Discharges Over Oil-Immersed Surfaces, The Propagation. Anderson, Liao.......218–26; disc. 226
Metals for Telephone Contacts, Roll-Welding Precious. Quinlan.......763–6
Meter, A Transducer-Type Frequency. Smith..497–500
Meter Standards, A New Test Board for Portable Watt-hour. Levitsky.......367–70; disc. 370
Meter, The Design of a New 2-Element Single Disk Polyphase. Schmidt.......513–16
Meters and Wattmeters, and a Novel Division Device, A Novel Circuit for Electronic Power Factor. Ahmed.......194–203; disc. 203
Method for Evaluation of Thermal Stability of Magnet Wire Enamel, A. Currin, Dexter.......227–31; disc. 231
Method of Making Screen Room Interference Measurements, A. Mortenson, Truax....746–50; disc. 750
Method of Solving Mathieu's Equation, A. Smith.......520–5
Minimizing and Mapping Sequential Circuits. Bennett.......443–7; disc. 796
Minimum Phase Network, What Is a. Balabanian, LePage.......785–7; disc. 787
Modulated Signals, Transfer Functions for Amplitude-. Hellerman.......729–31
Modulator as a Polarized Rectifier, The Ring. Hermont.......21–30
Modulus, Insulation Aging Characteristics as Measured by Shearing. Whitman, Scheideler.......232–8; disc. 238
Moisture Detector, A Continuous-Flow. Beavers, Shimanski, Timpane.......361–6; disc. 366
Moisture Equilibriums in Transformer Oil, Gas and. Kaufman, Shimanski, MacFadyen.......312–18; disc. 318
Momentum Transfer, A Polar Relay Using. Garbarino, Bisshopp.......731–3
Motors, Heat-Resistant Insulation Systems for. Herman, Mathes.......561–5
Motors, Magnetic Amplifier Control of D-C. Kusko, Nelson.......326–33; disc. 333
Movable Core Transformer as a Pickup for Industrial Telemetering Systems, The. Hornfeck, Wermelskirchen.......454–62
Multigrid Tube Circuit Design, Notes on the Use of Screen-to-Plate Transconductance in. Pullen.......648–51; disc. 651
Multiplexing of 900-Megacycle Radio Relay Systems, Coaxial Cavity Filters for. Kebby.......261–7
Multiplication with Transistors and Magnetic Cores, Four-Quadrant. Van Allen.......643–7; disc. 647
Multiplier, Analysis and Design of a Magnetic Frequency. Smith, Salihi.......99–106
Multistable Transistor Circuit, A. Henle.......568–71
Multistage Diode Logic Circuits, Engineering. Yokelson, Ulrich.......466–75
Multivibrator, A Variable Frequency Magnetic-Coupled. Van Allen.......356–61; disc. 361
Mutual-Inductance Transducers with High-Conductivity Reference Plates, Eddy-Current. Joseph, Newman.......39–49

N

Narrow-Band FM Receivers, Impulse Noise in. Lapin, Suran.......450–4
Network Analyzer Operation, Computer for Automatizing. Kusko, Heller.......208–13; disc. 213
Network, Engineering Problems and Performance Record of the Hawaiian Interisland Radio. Farinon, Walker.......571–6
Network of the Hawaiian Telephone Company, The Interisland Radio. Farinon.......529–36
Network, What Is a Minimum-Phase. Balabanian, LePage.......785–7; disc. 787
New Line of Direct-Acting Recording Instruments, A. Lunas, Nycz.......679–84; disc. 684
New Test Board for Portable Watt-hour Meter Standards. Levitsky.......367–70; disc. 370
New York State Thruway, A Communications System for the. DeWire, Evans.......253–61
Nickel-Iron to Magnetic Amplifier Performance, Relation of D-C Magnetic Properties of Oriented 48-Percent. Ward, Littmann.......422–7
900-Megacycle Radio Relay Systems, Coaxial Cavity Filters for Multiplexing of. Kebby.......261–7
Nitrogen, Electric Breakdown of Perfluorocarbon Vapors and Their Mixture With. Berberich, Works, Lindsay.......660–5; disc. 665
Noise in Narrow-Band FM Receivers, Impulse. Lapin, Suran.......450–4
Nonlinear Coupled Circuits, Analysis of. Part II. Ku.......439–43
Notes on the Use of Screen-to-Plate Transconductance in Multigrid Tube Circuit Design. Pullen.......648–51; disc. 651
Novel Circuit for Electronic Power Factor Meters and Wattmeters, and a Novel Division Device, A. Ahmed.......194–203; disc. 203
Nuclear Power Plant on an Integrated Electric System, Operation of a. Wilson.......751–4; disc. 754

O

Oil, Empirical Equations for the 60-Cycle Spark-over Gradients in Commercial. Gauster.......301–07; disc. 307
Oil, Gas and Moisture Equilibriums in Transformer. Kaufman, Shimanski, MacFadyen.......312–18; disc. 318
Oil-Immersed Surfaces, The Propagation Mechanism of Impulse Creepage Discharges Over. Anderson, Liao.......218–26; disc. 226

100-Megawatt Rectifier Installation, Field Tests on a. Dillard, Kiefer, Hague..........245–52; disc. 252
120-Degree Angles, Equations for the Inductance and Short-Circuit Forces of 3-Phase Busses Comprised of. Maier, Higgins..............334–8; disc. 338.
160,000-Ampere High Speed Magnetic Amplifier Design. Rosenstein..........................90–9
Open-Wire Carrier Systems in South Africa. Boyce.....................................272–81
Operation of a Nuclear Power Plant on an Integrated Electric System. Wilson..........751–4; disc. 754
Operation of Self-Saturating Magnetic Amplifiers, High-Frequency. Collins................500–05
Operational Magnetic Amplifier with Audio-Frequency Transistor Power Supply. Decker, Gourash..490–6
Optimum Core Configurations for Magnetic Amplifiers Using a Simplified Approach, An Analysis of. Bedford, Willis, Dodson....................62–70
Optimum Design of Cathode Followers, On the. Vallese.....................................5–8
Oscillator with Losses Associated with the Inductive Branch and Signal Applied in Grid Circuit, Analysis of a Triode. Mostafa.....................777–81
Oscillator with Low Second-Harmonic Distortion, A Symmetrical Transistor. Grim...767–70; disc. 770
Oscillographic Techniques for the Evaluation of Magnetic Amplifier Response. Critchlow.......607–10
Output Frequency Proportional to the D-C Input Voltage, A Switching Transistor D-C to A-C Converter Having an. Royer..............322–4; disc. 324
Output Stage, A New Full-Wave Magnetic Amplifier. Barnhart...............................139–41
Overload, Television Receiver Signal. Masucci..78–82

P

Perfluorocarbon Vapors and Their Mixture With Nitrogen, Electric Breakdown of. Berberich, Works, Lindsay..................660–5; disc. 665
Performance Magnetic Amplifiers in A-C Servo Applications, Transistor Demodulator for High. Decker.......................................121–3
Performance Record of the Hawaiian Interisland Radio Network, Engineering Problems and. Farinon, Walker.......................................571–6
Perpendicularly Superposed Magnetic Fields, The Application of. Heartz, Buelteman.......................................655–9; disc. 659
Phase Locking of Switching-Transistor Converters for Polyphase Power Supplies. Milnes........587–92
Phase Network, What Is a Minimum. Balabanian, LePage.......................785–7; disc. 787
Phenomena Accompanying Transient Low-Voltage Discharges in Liquid Dielectrics. Williams, Smith.......................................164–9
Photographic Light Source of High Intrinsic Brilliance, The Underwater Spark; A. Early, Martin.......................................788–90
Pickup Using Magnetic Fields, A Survey of Noncontacting Vibration. Clarke...............186–90
Plates, Eddy-Current Mutual-Inductance Transducers with High-Conductivity Reference. Joseph, Newman......................................39–49
Plotter, An Automatic Field. Gelfand, Shinn, Tuteur.......................................73–8
Poisson Distribution, Machine Testing for Deviation of Data from a. Hadden.......................155–7
Polar Relay Using Momentum Transfer, A. Garbarino, Bisshopp.............................731–3
Polarity Half Cycle Response Magnetic Amplifiers, Full Wave Reversible. House..................541–52
Polarization Synthesizer, An Elliptical. Friedman.......................................339–43; disc. 343
Polarized Rectifier, The Ring Modulator as a. Hermont.......................................21–30
Pole-Mounted A-C Operated Repeater for 12-Channel Open-Wire Carrier Systems, A. Searle, Schaefer.......................................791–5; disc. 795
Polyester Magnet Wire, Heat Resistant. Gilliam, Boldebuck, Elliott...........................557–60
Polyphase Meter, The Design of a New 2-Element Single Disk. Schmidt...........................513–16
Polyphase Power Supplies, Phase Locking of Switching-Transistor Converters for. Milnes.........587–92
Portable Telegraph Bias and Distortion Measuring Instrument A. Cannon.........401–06; disc. 406
Portable Watthour Meter Standards, A New Test Board for. Levitsky.................367–70; disc. 370
Position A-C Electromotive-Force Transducer and Its Application to Null-Type Recorders, An Angular. Guptill, McCarty.............................610–13
Potential, Measuring and Recording Atmospheric Electrostatic. Carroll, Hammond, Stewart.......................................517–20
Power Factor Meters and Wattmeters, and a Novel Division Device, A Novel Circuit for Electronic. Ahmed.......................194–203; disc. 203
Power Plant on an Integrated Electric System, Operation of a Nuclear. Wilson........751–4; disc. 754
Power Rectifiers, Properties of Silicon. Losco...106–11
Precision Torque Balance for the Measurement of Small Torques, A. Matthys.....................485–90
Presentation of Data on Broad-Band and Pulse Transformer Cores. Gillette, Oshima, Henderson, Rowe.......................................755–63
Printing, Electronic Computer for Color. Rose...268–72
Problems and Performance Record of the Hawaiian Interisland Radio Network, Engineering. Farinon Walker.......................................571–6
Problems in Film Pickup for Television Broadcasting, Basic. Gore.............................462–6
Progress in the Evaluation of Solid-Core High-Voltage Bushings. Brustle, Johnston, Scheideler.......................................387–91; disc. 391
Propagation Mechanism of Impulse Creepage Discharges Over Oil-Immersed Surfaces, The. Anderson, Liao.................218–26; disc. 226
Properties of Magnetic Fluids, Some. Coolidge, Halberg.......................................149–52
Properties of Silicon Power Rectifiers. Losco....106–11
Proportional Counters in X-Ray Diffraction, Use of. Laird, Zunick..............................734–7
Puerto Rico and the Virgin Islands, Very High Frequency Radio Link Between. McSweeny...781–5
Pulse Transformer Cores, Presentation of Data on Broad-Band and. Gillette, Oshima, Henderson, Rowe.......................................755–63
Pulse Transformer Service, A Tapered Strip Transmission Line for. Primozich, Schatz, Woodford.......................................158–61
Pulsers, Economic Design of Saturating Reactor Magnetic. Mathias, Williams..................169–71

Q

Quadrant Multiplication with Transistors and Magnetic Cores, Four-. Van Allen........643–7; disc. 647

R

Radiation on Dielectric Solids, The Effects of High-Energy Gamma. Klein, Mannal..........723–9
(Radio) A Communications System for the New York State Thruway. DeWire, Evans..........253–61
Radio Circuits, Equipment for Interisland. Walker, Nolan, Hoover..........................576–83
(Radio Communication) Impulse Noise in Narrow-Band FM Receivers. Lapin, Suran..........450–4
(Radio Communication) Presentation of Data on Broad-Band and Pulse Transformer Cores. Gillette, Oshima, Henderson, Rowe...........755–63
Radio Link Between Puerto Rico and the Virgin Islands, Very High Frequency. McSweeny.........781–5
Radio Network, Engineering Problems and Performance Record of the Hawaiian Interisland. Farinon, Walker.......................................571–6
Radio Network of the Hawaiian Telephone Company, The Inter-island. Farinon...............529–36
Radio Relay Systems, Coaxial Cavity Filters for Multiplexing of 900-Megacycle. Kebby.........261–7
Radio System, The Seattle-Victoria. Kistler....737–41
Railroad Signal and Communications Circuits, Aluminum Conductor Steel-Reinforced (ACSR) for. Adams.......................................627–9
Rapid Measurement of Impedance and Admittance. Salzberg, Marini...........................180–6
Ratio on Static and Dynamic Magnetic Properties of Toroidal Cores, Influence of ID/OD. Roberts, Van Nice.......................................599–607
Reactor Irradiation on Electrical Insulation, The Effect of. Figg, Bopp, Sisman, Robinson...717–23
Reactor Magnetic Pulsers, Economic Design of Saturating. Mathias, Williams...................169–71
Receiver Signal Overload, Television. Masucci...78–82
Receivers, Impulse Noise in Narrow-Band FM. Lapin, Suran.......................................450–4
Recognition of Possible Measurement Errors in D-C Dielectric Testing in the Field, The. Ross, Curdts.......................................630–5; disc. 635
Recommendations of IEC Technical Committee No. 24 on Electric and Magnetic Magnitudes and Units. Smith.......................................37–9
Recorders, An Angular Position A-C Electromotive-Force Transducer and Its Application to Null-Type. Guptill, McCarty................610–13
Recorders, Ticketfax Transmitters and. Zabriskie.......................................536–41
Recording Atmospheric Electrostatic Potential, Measuring and. Carroll, Hammond, Stewart.....517–20
Recording Instruments, A New Line of Direct-Acting. Lunas, Nyez..............679–84; disc. 684
Rectifier Installation, Field Tests on a 100-Megawatt. Dillard, Kiefer, Hague..........245–52; disc. 252
Rectifier, The Ring Modulator as a Polarized. Hermont.......................................21–30
Rectifiers for High Temperature Operation, Selenium. Isaacson.......................................552–7
Rectifiers, Properties of Silicon Power. Losco...106–11
Rectifiers, Regulator Curves and Transient Currents of Double-Way and Double-Wye. Jensen, Rettig.......................................281–7; disc. 287
Reference Plates, Eddy-Current Mutual-Inductance Transducers with High-Conductivity. Joseph, Newman.......................................39–49
Regulator Curves and Transient Currents of Double-Way and Double-Wye Rectifiers. Jensen, Rettig.......................................281–7; disc. 287
Relation of D-C Magnetic Properties of Oriented 48-Percent Nickel-Iron to Magnetic Amplifier Performance. Ward, Littmann...............422–7
Relay, Development of a Wire Contact. Markle...742–5
Relay Systems, Coaxial Cavity Filters for Multiplexing of 900-Megacycle Radio. Kebby..........261–7
Relay Using Momentum Transfer, A Polar. Garbarino, Bisshopp.............................731–3
Reliability Engineering, Science, Statistics and. Bear.......................................82–5
Repeater for 12-Channel Open-Wire Carrier Systems, A Pole-Mounted A-C Operated. Searle, Schaefer.......................................791–5; disc. 795
Report of Dielectric Tests on a Large Hydrogenerator. Duke, Ross, Johnson..............673–8; disc. 678
Resistance of Energized A-C Windings, A Circuit for Measuring the. Seely..........214–17; disc. 217
Resistor, A Sine Function. Nielsen, Roland.....447–50
Resonant Dielectric Amplifier, Voltage Gain of a. Sack, Penney...............................428–34
Response Magnetic Amplifiers, Full Wave Reversible Polarity Half Cycle. House..................541–52
Response, Oscillographic Techniques for the Evaluation of Magnetic Amplifier. Critchlow.........607–10
Ring Modulator as a Polarized Rectifier, The. Hermont.......................................21–30
Roll-Welding Precious Metals for Telephone Contacts. Quinlan.......................................763–6
Rotating Machines, Magnetic Flux Density Measurements in. Ley, Charmats...................525–8
Rural Subscriber Service, Telephone Lines for. Hochgraf, Watling............................171–6

S

Saturating Reactor Magnetic Pulsers, Economic Design of. Mathias, Williams...................169–71
Schematic Circuit Drawing, An Improved Detached Contact Type of. Meyer...................505–13
Science, Statistics, and Reliability Engineering. Bear.......................................82–5
Screen Room Interference Measurements, A Method of Making. Mortenson, Truax....746–50; disc. 750
Screen-to-Plate Transconductance in Multigrid Tube Circuit Design, Notes on the Use of. Pullen.......................................648–51; disc. 651
Seattle-Victoria Radio System, The. Kistler....737–41
Second-Harmonic Distortion, A Symmetrical Transistor Oscillator with Low. Grim.....767–70; disc. 770
Selenium Rectifiers for High Temperature Operation. Isaacson.......................................552–7
Self-Saturating Magnetic Amplifiers, High-Frequency Operation of. Collins...................500–05
Sequential Circuits, Minimizing and Mapping. Bennett.......................................443–7; disc. 796
Sequential Circuits of the Asymmetrical Type, Magnetic Amplifier. Fitzgerald...................685–91
Service, Telephone Lines for Rural Subscriber. Hochgraf, Watling............................171–6
Serviceboard Using Electronic Circuits, A New Telegraph. Davey, Hanley, Purvis.................30–7
Servo Amplifier, A Fast Response Magnetic. Kallander.......................................49–54
Servo Amplifiers, Elimination of Asymmetry Zero-Drift Errors in Magnetic. Geyger.......................................239–44; disc. 244
Servo Applications, Decicycle Magnetic Amplifier Systems for. Johnson, Rauch....667–72; disc. 672
Servo Applications, Transistor Demodulator for High Performance Magnetic Amplifiers in A-C. Decker.......................................121–3
Shearing Modulus, Insulation Aging Characteristics as Measured by. Whitman, Scheideler.......................................232–8; disc. 238
Sheet Insulation, Significant Factors in Thermal Aging Tests on Flexible. Dakin, Philofsky, Divens.......................................289–92; disc. 292
Shielding of Communication Cables. Gooding, Slade.......................................378–86; disc. 386
Shift Signaling Circuit for 45-Type Carrier Systems, Frequency-. Appert, Caruthers............613–19
Short-Circuit Forces of 3-Phase Busses Comprised of 120-Degree Angles, Equations for the Inductance and. Maier, Higgins..................334–8; disc. 338
Signal and Communications Circuits, Aluminum Conductor Steel-Reinforced (ACSR) for Railroad. Adams.......................................627–9
Signal Control, Textile Automation by. Casper.......................................176–80
Signal Overload, Television Receiver. Masucci...78–82

Signaling Circuit for 45-Type Carrier Systems, Frequency-Shift. Appert, Caruthers.........613–19
Signals, Transfer Functions for Amplitude-Modulated. Hellerman.........729–31
Significant Factors in Thermal Aging Tests on Flexible Sheet Insulation. Dakin, Philofsky, Divens.........289–92; disc. 292
Silicon Power Rectifiers, Properties of. Losco...106–11
Sine Function Resistor, A. Nielsen, Roland.....447–50
Single Disk Polyphase Meter, The Design of a New 2-Element. Schmidt.........513–16
60-Cycle Spark-over Gradients in Commercial Oil, Empirical Equations for the. Gauster.........301–07; disc. 307
Solid-Core High-Voltage Bushings, Progress in the Evaluation of. Brustle, Johnston, Scheideler.........387–91; disc. 391
Solids, The Effects of High-Energy Gamma Radiation on Dielectric. Klein, Mannal.........723–9
Some Electrical and Thermal Characteristics of Helium and Sulfur-Hexafluoride Mixtures. Cobine.........318–21; disc. 321
Some Properties of Magnetic Fluids. Coolidge, Halberg.........149–52
South Africa, Open-Wire Carrier Systems in. Boyce.........272–81
Spark; A Photographic Light Source of High Intrinsic Brilliance, The Underwater. Early, Martin.........788–90
Spark-over Gradients in Commercial Oil, Empirical Equations for the 60-Cycle. Gauster.........301–07; disc. 307
Stability of Magnet Wire Enamel, A Method for Evaluation of Thermal. Currin, Dexter.........227–31; disc. 231
Stage, A New Full-Wave Magnetic Amplifier Output. Barnhart.........139–41
Standards, A New Test Board for Portable Watthour Meter. Levitsky.........367–70; disc. 370
Standards for the Presentation of Data on Magnetic Amplifier Core Materials. (Committee Report).........598–9
Static and Dynamic Magnetic Properties of Toroidal Cores, Influence of ID/OD Ratio on. Roberts, Van Nice.........599–607
Static Magnetic Transistor Analogue Divider, A. Schaefer.........652–5; disc. 655
Static Time Delay Device Utilizing Transistors, A High Accuracy. Pittman.........54–8
Statistics and Reliability Engineering, Science. Bear.........82–5
Steel-Reinforced (ACSR) for Railroad Signal and Communications Circuits, Aluminum Conductor. Adams.........627–9
Stress on Magnetic Laminations, Effects of Transverse Compressional. Fischell.........190–3
Subscriber Service, Telephone Lines for Rural. Hochgraf, Watling.........171–6
Substitute for Mechanical Low-Level Choppers, Switching Transistors Used as a. Kruper.........141–4
Sulfur-Hexafluoride Mixtures, Some Electrical and Thermal Characteristics of Helium and. Cobine.........318–21; disc. 321
Superposed Magnetic Fields, The Application of Perpendicularly. Heartz, Buelteman.........655–9; disc. 659
Supplement to the Brune Synthesis, A. Reza....85–90
Surface and Volume Dielectric Losses. Chapman, Blickley, Szymkowiak.........343–8; disc. 348
Surface Types, Dielectric Failure of Volume and. Chapman, Frisco, Smith.........349–54; disc. 354
Survey of Noncontacting Vibration Pickup Using Magnetic Fields, A. Clarke.........186–90
Switches, Junction Transistors Used as. Bright...111–21
Switching-Transistor Converters for Polyphase Power Supplies, Phase Locking of. Milnes.........587–92
Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage, A. Royer.........322–4; disc. 324
Switching Transistors Used as a Substitute for Mechanical Low-Level Choppers. Kruper.....141–4
Symmetrical Transistor Oscillator with Low Second-Harmonic Distortion, A. Grim...767–70; disc. 770
Synthesis, A Supplement to the Brune. Reza....85–90
Synthesizer, An Elliptical Polarization. Friedman.........339–43; disc. 343
System, The Ticketfax. Hill.........583–6

T

Tachometer Indicator, Temperature Errors in a Dragmagnet Eddy-Current Disk Type of. Akeley, Fraizer.........418–22
Tapered Strip Transmission Line for Pulse Transformer Service, A. Primozich, Schatz, Woodford..158–61
(Telegraph) A Polar Relay Using Momentum Transfer. Garbarino, Bisshopp.........731–3
(Telegraph) An Envelope Delay-Measuring Instrument in the Audio-Frequency Range. Cannon...710–17
Telegraph Bias and Distortion Measuring Instrument, A Portable. Cannon.........401–06; disc. 406
Telegraph Carrier Terminal, A New Audio. Higgins, Gilbert.........58–62
Telegraph Facilities, A Fully Selective Telemetering System Employing. Smith, Forrest.........373–7; disc. 377
Telegraph Serviceboard Using Electronic Circuits, A New. Davey, Hanley, Purvis.........30–7
(Telegraph Systems) Textile Automation by Signal Control. Casper.........176–80
(Telegraph Systems) The Ticketfax System. Hill.........583–6
(Telegraph Systems) Ticketfax Electronics. Jelinek.........565–8
(Telegraph Systems) Ticketfax Transmitters and Records. Zabriskie.........536–41
Telemetering System Employing Telegraph Facilities, A Fully Selective. Smith, Forrest...373–7; disc. 377
Telemetering Systems, The Movable Core Transformer as a Pickup for Industrial. Hornfeck, Wermelskirchen.........454–62
Telephone Cable, A Transatlantic. Kelly, Radley, Gilman, Halsey.........124–36; disc. 136
Telephone Contacts, Roll-Welding Precious Metals for. Quinlan.........763–6
Telephone Lines for Rural Subscriber Service. Hochgraf, Watling.........171–6
Telephone Systems, Aluminum Die Castings for Carrier. Pedersen.........434–9
(Telephone) The Appraisal of Delays in Gate-Type Operation. Molnar.........475–85
Television Broadcasting, Basic Problems in Film Pickup for. Gore.........462–6
Television, Developments in Closed-Circuit. Kraus.........704–09
Television Receiver Signal Overload. Masucci...78–82
Temperature Errors in a Dragmagnet Eddy-Current Disk Type of Tachometer Indicator. Akeley, Fraizer.........418–22
Temperature Indicator, An Air-borne. Clark, Amey, Mergner.........1–5
Temperature Operation, Selenium Rectifiers for High. Isaacson.........552–7
Terminal, A New Audio Telegraph Carrier. Higgins, Gilbert.........58–62
Test Board for Portable Watthour Meter Standards, A New. Levitsky.........367–70; disc. 370
Testing for Deviation of Data from a Poisson Distribution, Machine. Hadden.........155–7
Testing in the Field, The Recognition of Possible Measurement Errors in D-C Dielectric. Ross, Curdts.........630–5; disc. 635
Tests on a Large Hydrogenerator, Report of Dielectric. Duke, Ross, Johnson.........637–8; disc. 678
Tests on a 100-Megawatt Rectifier Installation, Field. Dillard, Kiefer, Hague.........245–52; disc. 252
Textile Automation by Signal Control. Casper.........176–80
Theory of Magnetic Cross-Valves. Higa.........204–07; disc. 207
Thermal Aging Tests on Flexible Sheet Insulation, Significant Factors in. Dakin, Philofsky, Divens.........289–92; disc. 292
Thermal Characteristics of Helium and Sulfur-Hexafluoride Mixtures, Some Electrical and. Cobine.........318–21; disc. 321
Thermal Life of Enameled Wire by Laboratory Methods, Determination of. Sattler.........70–3
Thermal Stability of Magnet Wire Enamel, A Method for Evaluation of. Currin, Dexter.........227–31; disc. 231
Three-Phase Busses Comprised of 120-Degree Angles, Equations for the Inductance and Short-Circuit Forces of. Maier, Higgins.........334–8; disc. 338
Ticketfax Electronics. Jelinek.........565–8
Ticketfax System, The. Hill.........583–6
Ticketfax Transmitters and Recorders. Zabriskie.........536–41
Time Delay Device Utilizing Transistors, A High-Accuracy Static. Pittman.........54–8
Toroidal Cores, Influence of ID/OD Ratio on Static and Dynamic Magnetic Properties of. Roberts, Van Nice.........599–607
Torque Balance for the Measurement of Small Torques, A Precision. Matthys.........485–90
Tracer for Dynamic Characteristics of Magnetic Materials, An Accurate Electronic. Kittl.....407–18
Traffic Control, Data Transfer and Display Equipment for a Proposed System of Air. Fenimore.....145–9
Traffic Control Information, An Electronic System for Air. Kalb.........152–5
Transatlantic Telephone Cable, A. Kelly, Radley, Gilman, Halsey.........124–36; disc. 136
Transconductance in Multigrid Tube Circuit Design, Notes on the Use of Screen-to-Plate. Pullen.........648–51; disc. 651
Transducer and Its Application to Null-Type Recorders, An Angular Position A-C Electromotive-Force. Guptill, McCarty.........610–13
Transducer-Type Frequency Meter, A. Smith.........497–500
Transducers with High-Conductivity Reference Plates, Eddy-Current Mutual-Inductance. Joseph, Newman.........39–49
Transfer, A Polar Relay Using Momentum. Garbarino, Bisshopp.........731–3
Transfer Functions for Amplitude-Modulated Signals. Hellerman.........729–31
Transformer as a Pickup for Industrial Telemetering Systems, The Movable Core. Hornfeck, Wermelskirchen.........454–62
Transformer Cores, Presentation of Data on Broad-Band and Pulse. Gillette, Oshima, Henderson, Rowe.........755–63
Transformer Oil, Gas and Moisture Equilibriums in. Kaufman, Shimanski, MacFadyen.........312–18; disc. 318
Transformer Service, A Tapered Strip Transmission Line for Pulse. Primozich, Schatz, Woodford...158–61
Transient Currents of Double-Way and Double-Wye Rectifiers, Regulator Curves and. Jensen, Rettig.........281–7; disc. 287
Transient Low-Voltage Discharges in Liquid Dielectrics, Phenomena Accompanying. Williams, Smith.........164–9
Transistor Analogue Divider, A Static Magnetic. Schaefer.........562–5; disc. 655
Transistor Circuit, A Multistable. Henle......568–71
Transistor Converters for Polyphase Power Supplies, Phase Locking of Switching-. Milnes......587–92
Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage, A Switching. Royer.........322–4; disc. 324
Transistor Demodulator for High Performance Magnetic Amplifiers in A-C Servo Applications. Decker.........121–3
Transistor Oscillator with Low Second-Harmonic Distortion, A Symmetrical. Grim...767–70; disc. 770
Transistor Power Supply, Operational Magnetic Amplifier with Audio-Frequency. Decker, Gourash.........490–6
Transistors, A High-Accuracy Static Time Delay Device Utilizing. Pittman.........54–8
Transistors and Magnetic Cores, Four-Quadrant Multiplication with. Van Allen.......643–7; disc. 647
Transistors Used as a Substitute for Mechanical Low-Level Choppers, Switching. Kruper.......141–4
Transistors Used as Switches, Junction. Bright..111–21
Transmission Line for Pulse Transformer Service, A Tapered Strip. Primozich, Schatz, Woodford.........158–61
Transmitters and Recorders, Ticketfax. Zabriskie.........536–41
Transverse Compressional Stress on Magnetic Laminations, Effects of. Fischell.........190–3
Triode Oscillator with Losses Associated with the Inductive Branch and Signal Applied in Grid Circuit, Analysis of a. Mostafa.........777–81
Tube Circuit Design, Notes on the Use of Screen-to-Plate Transconductance in Multigrid. Pullen.........648–51; disc. 651
12-Channel Open-Wire Carrier Systems, A Pole-Mounted A-C Operated Repeater for. Searle, Schaefer.........791–5; disc. 795
24-Channel Cable Carrier and 12-Channel Open-Wire Carrier Systems, Combined Operation of. Searle, Schaefer.........770–6; disc. 777
Two-Element Single Disk Polyphase Meter, The Design of a New. Schmidt.........513–16

U

Underwater Spark; A Photographic Light Source of High Intrinsic Brilliance, The. Early, Martin.........788–90
Use of Proportional Counters in X-Ray Diffraction. Laird, Zunick.........734–7

V

Valves, Theory of Magnetic Cross-. Higa.........204–07; disc. 207
Vapors and Their Mixture With Nitrogen, Electric Breakdown of Perfluorocarbon. Berberich, Works, Lindsay.........660–5; disc. 665
Variable Frequency Magnetic-Coupled Multivibrator, A. Van Allen.........356–61; disc. 361
Vibration Pickup Using Magnetic Fields, A Survey of Noncontacting. Clarke.........186–90
Victoria Radio System, The Seattle-. Kistler...737–41
Virgin Islands, Very High Frequency Radio Link Between Puerto Rico and the. McSweeny.....781–5
Voltage, A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input. Royer.........322–4; disc. 324
Voltage Gain of a Resonant Dielectric Amplifier. Sack, Penney.........428–34

Vol.
Pt.
195

Volume and Surface Types, Dielectric Failure of. Chapman, Frisco, Smith........349–54; disc. 354
Volume Dielectric Losses, Surface and. Chapman, Blickley, Szymkowiak...........343–8; disc. 348

W

Watthour Meter Standards, A New Test Board for Portable. Levitsky..............367–70; disc. 370
Wattmeters, and a Novel Division Device, A Novel Circuit for Electronic Power Factor Meters and. Ahmed........................194–203; disc. 203
Welding Precious Metals for Telephone Contacts, Roll-. Quinlan...............................763–6
What Is a Minimum Phase Network? Balabanian, LePage...........................785–7; disc. 787
Windings, A Circuit for Measuring the Resistance of Energized A-C. Seely.........214–17; disc. 217
Wire by Laboratory Methods, Determination of Thermal Life of Enameled. Sattler....................70–3
Wire Carrier Systems, Combined Operation of 24-Channel Cable Carrier and 12-Channel Open-. Searle, Schaefer..................770–6; disc. 777
(Wire Communications) A Transatlantic Telephone Cable. Kelly, Radley, Gilman, Halsey...124–36; disc. 136
(Wire Communications) Aluminum Die Castings for Carrier Telephone Systems. Pedersen......434–9
(Wire Communications) Frequency-Shift Signaling Circuit for 45-Type Carrier Systems. Appert, Caruthers.................................613–19
(Wire Communications) Open-Wire Carrier Systems in South Africa. Boyce......................272–81
(Wire Communications) Telephone Lines for Rural Subscriber Service. Hochgraf, Watling.....171–5
Wire Contact Relay, Development of a. Markle..742–6
Wire Enamel, A Method for Evaluation of Thermal Stability of Magnet. Currin, Dexter.......................................227–31; disc. 231
Wire, Heat Resistant Polyester Magnet. Gilliam, Boldebuck, Elliott...........................557–60
Wire Insulation, Functional Evaluation of Magnetic. Brancato, Phillips................293–6; disc. 297

X

X-Ray Diffraction, Use of Proportional Counters in. Laird, Zunick.............................734–7

Z

Zero-Drift Errors in Magnetic Servo Amplifiers, Elimination of Asymmetry. Geyger..239–44; disc. 244

2. Author Index

Adams, H. W. Aluminum Conductor Steel-Reinforced (ACSR) for Railroad Signal and Communication Circuits. (55-559)........................627–9
Ahmed, M. Abdel-Halim. A Novel Circuit for Electronic Power Factor Meters and Wattmeters, and a Novel Division Device. (55-1)..194–203; disc. 204
Akeley, L. T.; J. J. Fraizer. Temperature Errors in a Dragmagnet Eddy-Current Disk Type of Tachmoeter Indicator. (55-242)................418–22
Amey, W. G.; W. R. Clark, G. C. Mergner. An Airborne Temperature Indicator. (55-2)........1–5
Anderson, J. G.; T. W. Liao. The Propagation Mechanism of Impulse Creepage Discharges Over Oil-Immersed Surfaces. (55-36)....218–26; disc. 226
Angwin, Sir Stanley. Disc.....................137
Appert, K. E.; R. S. Caruthers. Frequency-Shift Signaling Circuit for 45-Type Carrier Systems. (55-560)...............................613–19
Atkinson, R. W. Disc.........................355
Austin, G. F. Disc...........................371

B

Balabanian, Norman; W. P. LePage. What Is a Minimum Phase Network? (55-707)..785–7; disc. 787
Barnhart, P. W. A New Full-Wave Magnetic Amplifier Output Stage. (55-58)..................139–41
Bartlett, R. C. Disc..........................238
Bashe, C. J.; P. W. Jackson, H. A. Mussell, W. D. Winger. The Design of the IBM Type 702 System. (55-719).............................695–704
Bear, J. C. Science, Statistics, and Reliability Engineering. (55-151)........................82–5
Beavers, M. F.; E. J. Shimanski, E. F. Timpane. A Continuous-Flow Moisture Detector. (55-83)..361–6; disc. 366
Bedford, B. D.; C. H. Willis, G. C. Dodson. An Analysis of Optimum Core Configurations for Magnetic Amplifiers Using a Simplified Approach. (55-68)...62–70
Bennett, W. S. Minimizing and Mapping Sequential Circuits. (55-436)...............443–7; disc. 796
Berberich, L. J.; C. N. Works, E. W. Lindsay. Electric Breakdown of Perfluorocarbon Vapors and Their Mixture With Nitrogen. (55-493)...660–5; disc. 666
Bisshopp, K. E.; H. L. Garbarino. A Polar Relay Using Momentum Transfer. (55-175)......731–3
Blake, John T.; Schwenker, John E. Disc.........796
Blickley, L. F.; J. J. Chapman, E. A. Szymkowiak. Surface and Volume Dielectric Losses. (55-81)..343–8; disc. 348
Boldebuck, E. M.; W. F. Gilliam, J. R. Elliott. Heat Resistant Polyester Magnet Wire. (55-494)...557–60
Bopp, C. D.; J. C. Pigg, O. Sisman, C. C. Robinson. The Effect of Reactor Irradiation on Electrical Insulation. (55-694)......................717–23
Boyce, C. F. Open-Wire Carrier Systems in South Africa. (55-231)...........................272–81
Brancato, E. L.; R. S. Phillips. Functional Evaluation of Magnetic Wire Insulation. (55-48)...293–6; disc. 297
Bright, R. L. Junction Transistors Used as Switches. (55-156).................................111–21
Brustle, H. H.; D. L. Johnston, A. L. Scheideler. Progress in the Evaluation of Solid-Core High-Voltage Bushings. (55-176)......387–91; disc. 394
Buckley, Oliver E. Disc.........................136
Bueltemann, Herbert, Jr.; R. A. Heartz. The Application of Perpendicularly Superposed Magnetic Fields. (55-544)...............655–9; disc. 659
Burkhart, G. N., Jr. Disc.......................372

C

Cameron, A. W. W. Disc...............635, 665, 678
Camilli, G.; T. W. Liao, R. E. Plump. The Dielectric Behavior of Some Fluorogases and Their Mixtures. (55-440)......................637–42; disc. 642
Camilli, G.; R. E. Plump, T. W. Liao. Disc........665
Cannon, W. D. A Portable Telegraph Bias and Distortion Measuring Instrument. (55-166)...401–06; disc. 407
Cannon, W. D. An Envelope Delay-Measuring Instrument in the Audio-Frequency Range. (55-734)...710–17
Carroll, J. S.; S. B. Hammond, E. H. Stewart. Measuring and Recording Atmospheric Electrostatic Potential. (55-489)........................517–20
Caruthers, R. S.; K. E. Appert. Frequency-Shift Signaling Circuit for 45-Type Carrier Systems. (55-560)...............................613–19
Casper, Louis. Textile Automation by Signal Control. (55-174)...............................176–80
Charmatz, A.; B. J. Ley. Magnetic Flux Density Measurements in Rotating Machines. (55-491)...525–8
Chapman, J. J. Disc............................666
Chapman, J. J.; L. F. Blickley, E. A. Szymkowiak. Surface and Volume Dielectric Losses. (55-81)...343–8; disc. 348
Chapman, J. J.; L. J. Frisco, J. S. Smith. Dielectric Failure of Volume and Surface Types. (55-82)...349–54; disc. 355
Chen, K.; R. O. Decker. Disc..............647, 655
Clark, W. R.; W. G. Amey, G. C. Mergner. An Airborne Temperature Indicator. (55-2)........1–5
Clarke, Howard F. A Survey of Noncontacting Vibration Pickup Using Magnetic Fields. (55-168)...186–90
Cobine, J. D. Some Electrical and Thermal Characteristics of Helium and Sulfur-Hexafluoride Mixtures. (55-10)...............................318–21
Collins, H. William. High Frequency Operation of Self-Saturating Magnetic Amplifiers. (55-465)...500–05
Coolidge, J. E., Jr.; R. W. Halberg. Some Properties of Magnetic Fluids. (55-170)..............149–52
Critchlow, D. L. Oscillographic Techniques for the Evaluation of Magnetic Amplifier Response. (55-545)...............................607–10
Crites, W. R. Disc.............................287
Curdts, E. B.; C. W. Ross. The Recognition of Possible Measurement Errors in D-C Dielectric Testing in the Field. (55-456)............630–5; disc. 635
Currin, C. G.; J. F. Dexter. A Method for Evaluation of Thermal Stability of Magnet Wire Enamel. (55-46)......................227–31; disc. 232
Cypher, G. A. Disc.............................238

D

Dakin, T. W. Disc..............................642
Dakin, T. W.; H. M. Philofsky, W. C. Divens. Significant Factors in Thermal Aging Tests on Flexible Sheet Insulation. (55-37)......289–92; disc. 293
Davey, J. R.; F. H. Hanley, M. R. Purvis. A New Telegraph Serviceboard Using Electronic Circuits. (55-38).................................30–7
Davis, F. E. Disc...............................372
Decker, R. O. Transistor Demodulator for High Performance Magnetic Amplifiers in A-C Servo Applications. (55-67)......................121–3
Decker, R. O. Disc.............................672
Decker, R. O.; K. Chen. Disc..............647, 655
Decker, R. O.; F. Gourash. Operational Magnetic Amplifier with Audio-Frequency Transistor Power Supply. (55-463)........................490–6
DeWire, D. S.; H. A. Evans. A Communications System for the New York State Thruway. (55-201)...253–61
Dexter, J. F.; C. G. Currin. A Method for Evaluation of Thermal Stability of Magnet Wire Enamel. (55-46)......................227–31; disc. 232
Dillard, J. K.; J. Kiefer, C. S. Hague. Field Tests on a 100-Megawatt Rectifier Installation. (55-39).....................................245–52; disc. 252
Divens, W. C.; T. W. Dakin, H. M. Philofsky. Significant Factors in Thermal Aging Tests on Flexible Sheet Insulation. (55-37)......289–92; disc. 293
Doble, Frank C. Disc.................318, 366, 394
Dodson, G. C.; B. D. Bedford, C. H. Willis. An Analysis of Optimum Core Configurations for Magnetic Amplifiers Using a Simplified Approach. (55-68)................................62–70
Dortort, I. K. Disc.............................287
Douglas, John F. H. Disc.......................400
Duke, C. A. Disc...............................636
Duke, C. A.; C. W. Ross, J. S. Johnson. Report of Dielectric Tests on a Large Hydrogenerator. (55-236)......................673–8; disc. 678
Dunkle, W. F. Disc.............................393

E

Eager, G. S., Jr. Disc..........................355
Early, H. C.; E. A. Martin. The Underwater Spark; A Photographic Light Source of High Intrinsic Brilliance. (55-708)...................788–90
Elliott, J. R.; W. F. Gilliam, E. M. Boldebuck. Heat Resistant Polyester Magnet Wire. (55-494)..557–60
Evans, H. A.; D. S. DeWire. A Communications System for the New York State Thruway. (55-201)...253–61

F

Farinon, W. B. The Inter-island Radio Network of the Hawaiian Telephone Company. (55-534)..529–36
Farinon, W. B.; A. C. Walker. Engineering Problems and Performance Record of the Hawaiian Inter-island Radio Network. (55-535)...........571–6
Farr, H. K.; W. A. Keen, Jr. Improving Field Analogues Through Conformal Mapping. (55-105)...395–400; disc. 400
Fenimore, G. E. Data Transfer and Display Equipment for a Proposed System of Air Traffic Control. (55-149)................................145–9
Feth, G. C.; L. A. Finzi. Magnetic Amplifiers in Bistable Operation. (55-539)..............592–8
Finzi, L. A. Disc...............................333
Finzi, L. A.; G. C. Feth. Magnetic Amplifiers in Bistable Operation. (55-539)..............592–8
Fischell, R. E. Effects of Transverse Compressional Stress on Magnetic Laminations. (55-171)..190–3
Fitzgerald, Alan S. Magnetic Amplifier Sequential Circuits of the Asymmetrical Type. (55-620)...685–91
Foote, H. L. Cold-Cathode Counting Circuits. (55-197)................................161–4
Forrest, M. E., Jr.; C. W. Smith. A Fully Selective Telemetering System Employing Telegraph Facilities. (55-167)..................373–7; disc. 378
Fraizer, J. J.; L. T. Akeley. Temperature Errors in a Dragmagnet Eddy-Current Disk Type of Tachometer Indicator. (55-242)...............418–22
Fraser, W. Disc................................252
Fricke, R. F. Disc.............................217
Friedman, G. H. An Elliptical Polarization Synthesizer. (55-155)..................339–43; disc. 343
Frisco, L. J.; J. J. Chapman, J. S. Smith. Dielectric Failure of Volume and Surface Types. (55-82)...349–54; disc. 355

G

Garbarino, H. L.; K. E. Bisshopp. A Polar Relay Using Momentum Transfer. (55-175)......731–3
Gauster, W. F. Empirical Equations for the 60-Cycle Spark-over Gradients in Commercial Oil. (55-65)...301–07; disc. 309
Gelfand, R.; B. J. Shinn, F. B. Tuteur. An Automatic Field Plotter. (55-104)......................73–8
Geyger, W. A. Elimination of Asymmetry Zero-Drift Errors in Magnetic Servo Amplifiers. (55-59)...239–44; disc. 244
Gilbert, E. A.; C. A. Higgins. A New Audio Telegraph Carrier Terminal. (55-94).................58–62

ol. 74
. 1
955

Gilbert, J. J. Disc. 138
Gillette, P. R.; K. Oshima, K. W. Henderson, R. M. Rowe. Presentation of Data on Broad-Band and Pulse Transformer Cores. (55-750) 755–63
Gilliam, W. F.; E. M. Boldebuck, J. R. Elliott. Heat Resistant Polyester Magnet Wire. (55-494) 557–60
Gilman, G. W.; M. J. Kelly, G. Radley, R. J. Halsey. A Transatlantic Telephone Cable. (55-230) .... 124–36; disc. 138
Gooding, F. H.; H. B. Slade. Shielding of Communication Cables. (55-198) .... 378–86; disc. 387
Gordon, William T. Disc. 297
Gore, E. M. Basic Problems in Film Pickup for Television Broadcasting. (55-457) 462–6
Gourash, Francis; R. O. Decker. Operational Magnetic Amplifier with Audio-Frequency Transistor Power Supply. (55-463) 490–6
Gove, Harold E. Disc. 754
Greenwood, Allan N. Disc. 213
Grim, W. M., Jr. A Symmetrical Transistor Oscillator with Low Second-Harmonic Distortion. (55-740) .... 767–70; disc. 770
Guptill, E. C.; C. W. McCarty. An Angular Position A-C Electromotive-Force Transducer and Its Application to Null-Type Recorders. (55-547) 610–13

H

Hadden, F. A. Machine Testing for Deviation of Data from a Poisson Distribution. (55-154) 155–7
Hague, C. S.; J. K. Dillard, J. Kiefer. Field Tests on a 100-Megawatt Rectifier Installation. (55-39) .... 245–52; disc. 252
Halberg, R. W.; J. E. Coolidge, Jr. Some Properties of Magnetic Fluids. (55-170) 149–52
Halsey, R. J.; M. J. Kelly, G. Radley, G. W. Gilman. A Transatlantic Telephone Cable. (55-230) .... 124–36; disc. 138
Hammond, S. B.; J. S. Carroll, E. H. Stewart. Measuring and Recording Atmospheric Electrostatic Potential. (55-489) 517–20
Hanley, F. H.; J. R. Davey, M. R. Purvis. A New Telegraph Serviceboard Using Electronic Circuits. (55-38) 30–7
Heartz, Robert A.; H. Buelteman, Jr. The Application of Perpendicularly Superposed Magnetic Fields. (55-544) .... 655–9; disc. 659
Heller, P. N.; A. Kusko. Computer for Automatizing Network Analyzer Operation. (55-5) .... 208–13; disc. 213
Hellerman, H. Transfer Functions for Amplitude-Modulated Signals. (55-706) 729–31
Henderson, K. W.; P. R. Gillette, K. Oshima, R. M. Rowe. Presentation of Data on Broad-Band and Pulse Transformer Cores. (55-750) 755–63
Henle, R. A. A Multistable Transistor Circuit. (55-497) 568–71
Herman, H. J.; K. N. Mathes. Heat-Resistant Insulation Systems for Motors. (55-495) 561–5
Hermont, A. J. The Ring Modulator as a Polarized Rectifier. (55-14) 21–30
Higa, W. H. Theory of Magnetic Cross-Valves. (55-8) .... 204–07; disc. 208
Higgins, C. A.; E. A. Gilbert. A New Audio Telegraph Carrier Terminal. (55-94) 58–62
Higgins, T. J. Disc. 787
Higgins, T. J.; J. W. Maier. Equations for the Inductance and Short-Circuit Forces of 3-Phase Busses Comprised of 120-Degree Angles. (55-126) .... 334–8; disc. 338
Hill, A. S. The Ticketfax System. (55-501) 583–6
Hochgraf, Lester; R. G. Watling. Telephone Lines for Rural Subscriber Service. (55-229) 171–6
Holtz, F. C. Disc. 371
Hooper, E. T.; J. J. Suozzi. An All-Magnetic Audio Amplifier System. (55-70) .... 297–300; disc. 300
Hoover, V. P.; A. C. Walker, E. E. Nolan. Equipment for Interisland Radio Circuits. (55-536) 576–83
Hornfeck, A. J.; L. M. Wermelskirchen. The Movable Core Transformer as a Pickup for Industrial Telemetering Systems. (55-438) 454–62
House, C. B. Full-Wave Reversible-Polarity Half-Cycle-Response Magnetic Amplifiers. (55-537) 541–52

I

Isaacson, G. O. Selenium Rectifiers for High-Temperature Operation. (55-518) 552–7

J

Jackson, D. C., Jr. Disc. 651
Jackson, P. W.; C. J. Bashe, H. A. Mussell, W. D. Winger. The Design of the IBM Type 702 System. (55-719) 695–704
Jackson, R. R. Disc. 659
Jelinek, C., Jr. Ticketfax Electronics. (55–502) 565–8
Jensen, L. E.; C. E. Rettig. Regulator Curves and Transient Currents of Double-Way and Double-Wye Rectifiers. (55-45) .... 281–7; disc. 288
Johnson, D. L.; H. H. Brustle, A. L. Scheideler. Progress in the Evaluation of Solid-Core High-Voltage Bushings. (55-176) .... 387–91; disc. 394
Johnson, J. S. Disc. 636
Johnson, J. S.; C. A. Duke, C. W. Ross. Report of Dielectric Tests on a Large Hydrogenerator. (55-236) .... 673–8; disc. 678
Johnson, Leo J.; S. E. Rauch. Decicycle Magnetic Amplifier Systems for Servo Applications. (55-542) .... 667–72; disc. 672
Joseph, Horace M.; N. Newman. Eddy-Current Mutual-Inductance Transducers With High-Conductivity Reference Plates. (55-4) 39–49

K

Kalb, R. M. An Electronic System for Air Traffic Control Information. (55-147) 152–5
Kallander, J. W. A Fast Response Magnetic Servo Amplifier. (55-57) 49–54
Kaufman, R. B.; E. J. Shimanski, K. W. MacFadyen. Gas and Moisture Equilibriums in Transformer Oil. (55-66) .... 312–18; disc. 318
Kebby, M. H. Coaxial Cavity Filters for Multiplexing of 900-Megacycle Radio Relay Systems. (55-200) 261–7
Keen, W. A., Jr.; H. K. Farr. Improving Field Analogues Through Conformal Mapping. (55-105) .... 395–400; disc. 400
Kelly, Mervin J.; G. Radley, G. W. Gilman, R. J. Halsey. A Transatlantic Telephone Cable. (55-230) .... 124–36; disc. 138
Kiefer, John; J. K. Dillard, C. S. Hague. Field Tests on a 100-Megawatt Rectifier Installation. (55-39) .... 245–52; disc. 252
Kistler, R. E. The Seattle-Victoria Radio System. (55-752) 737–41
Kittl, Emil. An Accurate Electronic Tracer for Dynamic Characteristics of Magnetic Materials. (55-454) 407–18
Klein, Philipp; C. Mannal. The Effects of High-Energy Gamma Radiation on Dielectric Solids. (55-741) 723–9
Knopp, Harold L. Disc. 372
Kraus, Max H. Developments in Closed-Circuit Television. (55-687) 704–09
Kruper, A. P. Switching Transistors Used as a Substitute for Mechanical Low-Level Choppers. (55-157) 141–4
Ku, Y. H. Analysis of Nonlinear Coupled Circuits. Part II. (55-428) 439–43
Kusko, A.; P. N. Heller. Computer for Automatizing Network Analyzer Operation. (55-5) .... 208–13; disc. 213
Kusko, A.; J. G. Nelson. Magnetic Amplifier Control of D-C Motors. (55-60) .... 326–33; disc. 333

L

Laird, H. R.; M. J. Zunick. Use of Proportional Counters in X-Ray Diffraction. (55-739) 734–7
Langlois, C. A. Disc. 252
Lanigan, L. P. Disc. 338
Lapin, S. P.; J. J. Suran. Impulse Noise in Narrow-Band FM Receivers. (55-429) 450–4
Lehtonen, E. Disc. 300
Leon, H. I.; A. B. Rosenstein. Inductive Load Instability in Magnetic Amplifiers. (55-21) 8–20
LePage, Wilbur R.; N. Balabanian. What Is a Minimum Phase Network? (55-707) .... 785–7; disc. 787
Levitsky, F. J. A New Test Board for Portable Watthour Meter Standards. (55-150) .... 267–70; disc. 373
Ley, B. J.; A. Charmatz. Magnetic Flux Density Measurements in Rotating Machines. (55-491) 525–8
Liao, T. W.; J. G. Anderson. The Propagation Mechanism of Impulse Creepage Discharges Over Oil-Immersed Surfaces. (55-36) .... 218–26; disc. 226
Liao, T. W.; G. Camilli, R. E. Plump. The Dielectric Behavior of Some Fluorogases and Their Mixtures. (55-440) .... 637–42; disc. 642
Liao, T. W.; G. Camilli, R. E. Plump. Disc. 665
Lindsay, E. W.; L. J. Berberich, C. N. Works. Electric Breakdown of Perfluorocarbon Vapors and Their Mixture With Nitrogen. (55-493) 660–5; disc. 666
Littmann, M. F.; C. E. Ward. Relation of D-C Magnetic Properties of Oriented 48-Percent Nickel-Iron to Magnetic Amplifier Performance. (55-455) 422–7
Losco, E. F. Properties of Silicon Power Rectifiers. (55-152) 106–11
Lunas, L. J.; J. C. Nycz. A New Line of Direct-Acting Recording Instruments. (55-484) 679–84; disc. 685

M

MacFadyen, K. W.; R. B. Kaufman, E. J. Shimanski. Gas and Moisture Equilibriums in Transformer Oil. (55-66) .... 312–18; disc. 318
Maier, J. W.; T. J. Higgins. Equations for the Inductance and Short-Circuit Forces of 3-Phase Busses Comprised of 120-Degree Angles. (55-126) .... 334–8; disc. 338
Mannal, Clifford; P. H. Klein. The Effects of High-Energy Gamma Radiation on Dielectric Solids. (55-741) 723–9
Manning, M. L. Disc. 292
Marini, J. W.; B. Salzberg. Rapid Measurement of Impedance and Admittance. (55-211) 180–6
Markle, R. E. Development of a Wire Contact Relay. (55-763) 742–5
Martin, E. A.; H. C. Early. The Underwater Spark; A Photographic Light Source of High Intrinsic Brilliance. (55-708) 788–90
Masucci, C. Television Receiver Signal Overload. (55-139) 78–82
Mathes, K. N.; C. J. Herman. Heat-Resistant Insulation Systems for Motors. (55-495) 561–5
Mathias, R. A.; E. M. Williams. Economic Design of Saturating Reactor Magnetic Pulsers. (55-179) 169–71
Matthys, R. J. A Precision Torque Balance for the Measurement of Small Torques. (55-486) 485–90
McCarty, C. W.; E. C. Guptill. An Angular Position A-C Electromotive-Force Transducer and Its Application to Null-Type Recorders. (55-547) 610–13
McCreary, Harold J. Disc. 207
McSweeny, Roger. Very High Frequency Radio Link Between Puerto Rico and the Virgin Islands. (55-751) 781–5
Meahl, H. R. Disc. 386, 750
Mergner, G. C.; W. R. Clark, W. G. Amey. An Airborne Temperature Indicator. (55-2) 1–5
Meyer, F. T. An Improved Detached-Contact-Type of Schematic Circuit Drawing. (55-467) 505–13
Milnes, A. G. Phase Locking of Switching-Transistor Converters for Polyphase Power Supplies. (55-543) 587–92
Milnes, A. G. Disc. 324
Molnar, Imre. The Appraisal of Delays in Gate-Type Operation. (55-466) 475–85
Morris, J. A. Disc. 684
Mortenson, K. E.; C. J. Truax. A Method of Making Screen Room Interference Measurements. (55-737) .... 746–50; disc. 750
Morton, Lysle W. Disc. 252
Moses, Graham Lee. Disc. 636, 678
Mostafa, Abd El-Samie. Analysis of a Triode Oscillator with Losses Associated with the Inductive Branch and Signal Applied in Grid Circuit. (55-705) 777–81
Mussell, H. A.; C. J. Bashe, P. W. Jackson, W. D. Winger. The Design of the IBM Type 702 System. (55-719) 695–704

N

Nelson, J. G.; A. Kusko. Magnetic Amplifier Control of D-C Motors. (55-60) .... 326–33; disc. 333
Newman, Nathan; H. M. Joseph. Eddy-Current Mutual-Inductance Transducers With High-Conductivity Reference Plates. (55-4) 39–49
Nielsen, K. L.; E. H. Roland. A Sine Function Resistor. (55-430) 447–50
Nolan, E. E.; A. C. Walker, V. P. Hoover. Equipment for Interisland Radio Circuits. (55-536) 576–83
Nycz, J. C.; L. J. Lunas. A New Line of Direct-Acting Recording Instruments. (55-484) 679–84; disc. 685

O

Oliver, F. S. Disc. 635
Oshima, Kazu; P. R. Gillette, K. W. Henderson, R. M. Rowe. Presentation of Data on Broad-Band and Pulse Transformer Cores. (55-750) 755–63

P

Parker, John C. Disc. 636
Patton, Henry W. Disc. 244
Pedersen, Ludwig. Aluminum Die Castings for Carrier Telephone Systems. (55-259) 434–9
Penney, G. W.; E. A. Sack. Voltage Gain of a Resonant Dielectric Amplifier. (55-71) 428–34
Petri, J. N. Disc. 777, 795
Phillips, R. S.; E. L. Brancato. Functional Evaluation of Magnetic Wire Insulation. (55-48) .... 293–6; disc. 297
Philofsky, H. M.; T. W. Dakin, W. C. Divens. Significant Factors in Thermal Aging Tests on Flexible Sheet Insulation. (55-37) .... 289–92; disc. 293

Vol
Pt.
195

Pigg, J. C.; C. D. Bopp, O. Sisman, C. C. Robinson. The Effect of Reactor Irradiation on Electrical Insulation. (55-694)....717–23

Pittman, G. F., Jr. A High-Accuracy Static Time Delay Device Utilizing Transistors. (55-74)..54–8

Plump, R. E. Disc....321

Plump, R. E.; G. Camilli, T. W. Liao. The Dielectric Behavior of Some Fluorogases and Their Mixtures. (55-440)....637–42; disc. 642

Plump, R. E.; G. Camilli, T. W. Liao. Disc.....665

Primozich, F. G.; E. R. Schatz, J. B. Woodford. A Tapered Strip Transmission Line for Pulse Transformer Service. (55-177)....158–61

Pritchett, W. S. Disc....371

Pullen, K. A., Jr. Notes on the Use of Screen-to-Plate Transconductance in Multigrid Tube Circuit Design. (55-500)....548–51; disc. 651

Purvis, M. R.; J. R. Davey, F. H. Hanley. A New Telegraph Serviceboard Using Electronic Circuits. (55-38)....30–7

Q

Quinlan, A. L. Roll-Welding Precious Metals for Telephone Contacts. (55-762)....763–6

R

Radley, Sir Gordon; M. J. Kelly, G. W. Gilman, R. J. Halsey. A Transatlantic Telephone Cable. (55-230)....124–36; disc. 138

Rauch, S. E.; L. J. Johnson. Decicycle Magnetic Amplifier Systems for Servo Applications. (55-542)....667–72; disc. 672

Rettig, C. E.; L. E. Jensen. Regulator Curves and Transient Currents of Double-Way and Double-Wye Rectifiers. (55-45)....281–7; disc. 288

Reza, F. M. A Supplement to the Brune Synthesis. (55-140)....85–90

Roberts, R. W.; R. I. Van Nice. Influence of ID/OD Ratio on Static and Dynamic Magnetic Properties of Toroidal Cores. (55-538)....599–607

Robinson, C. C.; J. C. Pigg, C. D. Bopp, O. Sisman. The Effect of Reactor Irradiation on Electrical Insulation. (55-694)....717–23

Roland, E. H.; K. L. Nielsen. A Sine Function Resistor. (55-430)....447–50

Rose, H. E. Electronic Computer for Color Printing. (55-180)....268–72

Rosenberger, H. J. Disc....137

Rosenstein, Allen B. 160,000-Ampere High Speed Magnetic Amplifier Design. (55-61)....90–9

Rosenstein, Allen B.; H. I. Leon. Inductive Load Instability in Magnetic Amplifiers. (55-21)....8–20

Ross, C. W.; E. B. Curdts. The Recognition of Possible Measurement Errors in D-C Dielectric Testing in the Field. (55-456)....630–5; disc. 635

Ross, C. W.; C. A. Duke, J. S. Johnson. Report of Dielectric Tests on a Large Hydrogenerator. (55-236)....673–8; disc. 678

Rowe, R. M.; P. R. Gillette, K. Oshima, K. W. Henderson. Presentation of Data on Broad-Band and Pulse Transformer Cores. (55-750)....755–63

Royer, G. H. A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage. (55-73)....322–4; disc. 325

Rufleth, W. E. Disc....377

S

Sack, E. A.; G. W. Penney. Voltage Gain of a Resonant Dielectric Amplifier. (55-71)....428–34

Salihi, Jalal T.; O. J. M. Smith. Analysis and Design of a Magnetic Frequency Multiplier. (55-69)....99–106

Salzberg, B.; J. W. Martini. Rapid Measurement of Impedance and Admittance. (55-211)....180–6

Sandell, Donald H. Disc....338

Sattler, F. A. Determination of Thermal Life of Enameled Wire by Laboratory Methods. (55-47)....70–3

Schaefer, D. H. A Static Magnetic Transistor Analogue Divider. (55-451)....652–5; disc. 655

Schaefer, D. H. Formulation of Magnetic Amplifier Circuit Equations. (55-546)....619–27

Schaefer, R. A.; G. W. Searle. A Pole-Mounted A-C Operated Repeater for 12-Channel Open-Wire Carrier Systems. (55-682)....791–5

Schaefer, R. A.; G. W. Searle. Combined Operation of 24-Channel Cable Carrier and 12-Channel Open-Wire Carrier Systems. (55-683)....770–6; disc. 777

Scharff, Samuel A. Disc....377

Schatz, E. R.; F. G. Primozich, J. B. Woodford. A Tapered Strip Transmission Line for Pulse Transformer Service. (55-177)....158–61

Scheideler, A. L.; H. H. Brustle, D. L. Johnston. Progress in the Evaluation of Solid-Core High-Voltage Bushings. (55-176).....387–91; disc. 394

Scheideler, A. L.; L. C. Whitman. Insulation Aging Characteristics as Measured by Shearing Modulus. (55-35)....232–8; disc. 238

Schmidt, W. J. The Design of a New 2-Element Single Disk Polyphase Meter. (55-483)....513–16

Schwenker, John E.; Blake, John T. Disc....796

Searle, G. W.; R. A. Schaefer. A Pole-Mounted A-C Operated Repeater for 12-Channel Open-Wire Carrier Systems. (55-682)....791–5

Searle, G. W.; R. A. Schaefer. Combined Operation of 24-Channel Cable Carrier and 12-Channel Open-Wire Carrier Systems. (55-683)..770–6; disc. 777

Seely, R. E. A Circuit for Measuring the Resistance of Energized A-C Windings. (55-23)214–17; disc. 218

Shea, R. F. Disc....770

Shimanski, E. J.; M. F. Beavers, E. F. Timpane. A Continuous-Flow Moisture Detector. (55-83)....361–6; disc. 366

Shimanski, E. J.; R. B. Kaufman, K. W. MacFadyen. Gas and Moisture Equilibriums in Transformer Oil. (55-66)....312–18; disc. 318

Shinn, B. J.; R. Gelfand, F. B. Tuteur. An Automatic Field Plotter. (55-104)....73–8

Sieffert, L. E.; H. P. Walker. Disc....297

Siegel, C. M. Disc....338

Sisman, O.; J. C. Pigg, C. D. Bopp, C. C. Robinson. The Effect of Reactor Irradiation on Electrical Insulation. (55-694)....717–23

Slade, H. B.; F. H. Gooding. Shielding of Communication Cables. (55-198)....378–86; disc. 387

Smith, C. W.; M. E. Forrest, Jr. A Fully Selective Telemetering System Employing Telegraph Facilities. (55-167)....373–7; disc. 378

Smith, J. J. A Method of Solving Mathieu's Equation. (55-490)....520–5

Smith, J. J. Recommendations of IEC Technical Committee No. 24 on Electric and Magnetic Magnitudes and Units. (55-49)....37–9

Smith, J. S.; J. J. Chapman, L. J. Frisco. Dielectric Failure of Volume and Surface Types. (55-82)....349–54; disc. 355

Smith, Otto J. M.; J. T. Salihi. Analysis and Design of a Magnetic Frequency Multiplier. (55-69).99–106

Smith, R. E.; E. M. Williams. Phenomena Accompanying Transient Low-Voltage Discharges in Liquid Dielectrics....164–9

Smith, U. L. A Transducer-Type Frequency Meter. (55-485)....497–500

Sommerman, G. M. L. Disc. 226, 231, 348, 354, 366, ...392

Spencer, J. H. Disc....370

Stewart, E. H.; J. S. Carroll, S. B. Hammond. Measuring and Recording Atmospheric Electrostatic Potential. (55-489)....517–20

Suozzi, J. J.; E. T. Hooper. An All-Magnetic Audio Amplifier System. (55-70).....297–300; disc. 300

Suran, J. J.; S. P. Lapin. Impulse Noise in Narrow-Band FM Receivers. (55-429)....450–4

Szymkowiak, E. A.; J. J. Chapman, L. F. Blickley. Surface and Volume Dielectric Losses. (55-81)....343–8; disc. 348

T

Terry, V. J. Disc....401

Timpane, E. F.; M. F. Beavers, E. J. Shimanski. A Continuous-Flow Moisture Detector. (55-83)....361–6; disc. 366

Truax, C. J.; K. E. Mortenson. A Method of Making Screen Room Interference Measurements. (55-737)....746–50; disc. 750

Tuteur, F. B.; R. Gelfand, B. J. Shinn. An Automatic Field Plotter. (55-104)....73–8

U

Uchrin, George C. Disc....324, 361

V

Vallese, L. M. On the Equivalent Circuits of Linear Amplifiers. (55-665)....691–4

Vallese, L. M. On the Optimum Design of Cathode Followers. (55-7)....5–8

Van Allen, R. L. A Variable Frequency Magnetic Coupled Multivibrator. (55-75) 356–61; disc. 361

Van Allen, R. L. Four-Quadrant Multiplication with Transistors and Magnetic Cores. (55-461)....643–7; disc. 647

Van Ness, James E. Disc....203, 213

Van Nice, R. I.; R. W. Roberts. Influence of ID/OD Ratio on Static and Dynamic Magnetic Properties of Toroidal Cores. (55-538)....599–607

W

Walker, A. C.; W. B. Farinon. Engineering Problems and Performance Record of the Hawaiian Interisland Radio Network. (55-535)....571–6

Walker, A. C.; E. E. Nolan, V. P. Hoover. Equipment for Interisland Radio Circuits. (55-536)...576–83

Walker, H. P.; L. E. Sieffert. Disc....297

Ward, C. E.; M. F. Littmann. Relation of D-C Magnetic Properties of Oriented 48-Percent Nickel-Iron to Magnetic Amplifier Performance. (55-455)....422–7

Watling, R. G.; L. Hochgraf. Telephone Lines for Rural Subscriber Service. (55-229)....171–6

Watts, J. M. Disc....343

Wermelskirchen, L. M.; A. J. Hornfeck. The Movable Core Transformer as a Pickup for Industrial Telemetering Systems. (55-438)....454–62

Whitman, L. C.; A. L. Scheideler. Insulation Aging Characteristics as Measured by Shearing Modulus. (55-35)....232–8; disc. 238

Wickham, W. H. Disc....391

Williams, E. M.; R. A. Mathias. Economic Design of Saturating Reactor Magnetic Pulsers. (55-179)....169–71

Williams, E. M.; R. E. Smith. Phenomena Accompanying Transient Low-Voltage Discharges in Liquid Dielectrics....164–9

Willis, C. H.; B. D. Bedford, G. C. Dodson. An Analysis of Optimum Core Configurations for Magnetic Amplifiers Using a Simplified Approach. (55-68)....62–70

Wilson, N. E. Operation of a Nuclear Power Plant on an Integrated Electric System. (55-749)....751–4; disc. 754

Wilson, W. R. Disc....307

Winger, W. D.; C. J. Bashe, P. W. Jackson, H. A. Mussell. The Design of the IBM Type 702 System. (55-719)....695–704

Woodford, J. B.; F. G. Primozich, E. R. Schatz. A Tapered Strip Transmission Line for Pulse Transformer Service. (55-177)....158–61

Works, C. N. Disc....354, 642

Works, C. N.; L. J. Berberich, E. W. Lindsay. Electric Breakdown of Perfluorocarbon Vapors and Their Mixture With Nitrogen. (55-493)....660–5; disc. 666

Wyatt, Kenneth S. Disc....137, 665

Y

Yokelson, B. J.; W. Ulrich. Engineering Multistage Diode Logic Circuits. (55-431)....466–75

Z

Zabriskie, D. M. Ticketfax Transmitters and Recorders. (55-503)....536–41

Zpaf, Thomas L. Disc....372

Zunick, M. J.; H. R. Laird. Use of Proportional Counters in X-Ray Diffraction. (55-739)...7347–

An Air-borne Temperature Indicator

W. R. CLARK
MEMBER AIEE

W. G. AMEY
MEMBER AIEE

G. C. MERGNER
ASSOCIATE MEMBER AIEE

THIS paper describes an air-borne temperature indicator which was developed specifically for the U. S. Army Signal Corps, Belmar, N. J. While the specification indicated that the instrument should be of the electromechanical null-balance type,[1] many of the required features represented significant departures from those of somewhat similar instruments used in industrial applications. Operation from a 115-volt (±10 per cent) 400-cycle (±5 per cent) power source was required, and the instrument was to have a range of −70 to +60 degrees centigrade (C) (4.92 millivolts for a copper-constantan thermocouple). The full-scale response time was specified to be less than 10 seconds, the minimum sensitivity 4 microvolts (0.14 C), and the limit of error ±1 C. Operation in accordance with these specifications, as well as storage without power connected, over an ambient temperature range of −55 to +55 C was required; this requirement precluded the use of dry batteries or standard cells. The indicator had to be small in size and weight, be sufficiently rugged to withstand severe shock and vibration, and develop sufficient torque to position a synchro transmitter.

These specifications indicated that there were two principal problems to be solved:

1. The development of a compact, lightweight, electromechanical balancing system which would have the required sensitivity, torque, response time, and freedom from adverse effects caused by shock or vibration.

2. The development of a new potentiometric system which would ensure the desired accuracy over the entire range of ambient temperature without employing batteries or standard cells.

General Description

The instrument developed meets the foregoing specifications. It is housed in a welded aluminum case designed for panel mounting and gasketed to be dust-proof. The main scale indications can be seen through a window, as shown in Fig. 1. The circular scale is marked with fluorescent paint in 1-degree divisions from −70 to +60 C. All controls necessary to the normal operation of the indicator and connectors for external circuits are mounted on the aluminum front panel which is rigidly attached to the aluminum main frame. This assembly slides in and out of the case as shown in Fig. 2. On the frame are mounted the balancing motor and synchro transmitter. On the back of the frame, Fig. 3, are mounted components associated with the potentiometer circuit and its current supply: measuring slide wire with its associated resistance coils and rheostats, compensating coils, voltage reference tubes, and thermostat. The balancing amplifier and the electronic portion of the regulated potentiometer current supply are mounted on an aluminum chassis, Fig. 4, connections to which are made through a multipin connector.

Balancing System

The balancing system, shown schematically in Fig. 5, is similar in principle to that used in the present-day Leeds & Northrup type-*G* recorders, except that the synchronous converter, balancing amplifier, and motor had to be specially designed to operate from a 400-cycle power source. A special 400-cycle synchronous converter was developed to meet the stringent requirements of this application. Tests of commercially available converters indicated insufficient long-term stability, excessive pickup from the drive coil, excessive contact noise, and susceptibility to ambient temperature changes. Operation of the new converter is independent of normal line voltage or line frequency changes. The drive coil requires 2.4 watts at 40 volts. Magnetic and electrostatic pickup in the low-level circuit of the converter is negligible at the operating level of the indicator. The converter does not introduce any appreciable noise into the amplifier and does not produce any significant zero offset of the indicator.

The balancing amplifier must have a sensitivity of approximately 4 microvolts (0.14 C). The unusually wide range of ambient temperature requires special precautions throughout the entire low-level circuit in order to prevent generation of voltages from thermal causes. The single-section resistance-capacitance filter used in the input circuit attenuates pickup in the thermocouple leads and ensures proper damping of the indicator without oscillation. The capacitor developed specially for this filter is relatively free from self-generated spurious currents over the entire range of ambient temperature. The synchronous converter converts any difference between the potentiometer voltage and the thermocouple voltage to a 400-cycle voltage which is stepped up by the input transformer and amplified by a 4-stage voltage amplifier (one *12AY7* twin triode and one *12AX7* twin triode). This part of the amplifier is enclosed in a shielded compartment. All leads enter via feed-through capacitors which filter out external, transient, high-frequency disturbances. The last stage of the voltage amplifier drives a single tube *6AQ5* power output stage which in turn drives the balancing motor. This tube supplies 3 watts at 400 cycles to the motor. A single rectifier tube supplies the direct-current voltage for both the amplifier and the regulated potentiometer supply. Because of the wide range of ambient temperatures no electrolytic capacitors are used in the indicator.

All heaters are operated from alternating current, but the input tube must be of a type which is inherently free from noise and hum. The level at which this tube operates requires using a special shielded socket to shield the grid and plate pins from the alternating current of the heater leads.

The motor is a 2-phase unit which delivers a maximum torque of 1 ounce-inch. It has a built-in gear reduction of

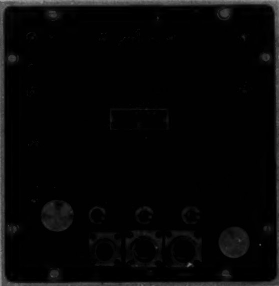
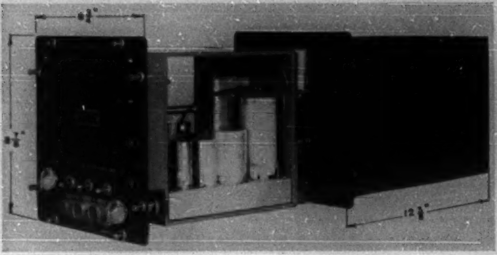


Fig. 1. Air-borne temperature indicator, front view

Paper 55-2, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted March 24, 1954; made available for printing October 27, 1954.

W. R. Clark, W. G. Amey, and G. C. Mergner are with the Leeds & Northrup Company, Philadelphia, Pa.

The 400-cycle converter used in the indicator was developed by W. McAdam and R. E. Tarpley.



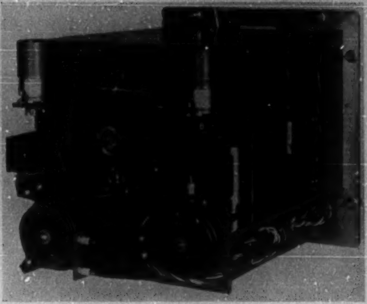


Fig. 2 (above). Indicator, case withdrawn

Fig. 3 (right). Slide wire and associated components

30 to 1 and a single external reduction of 16 to 1 to couple the motor to the slide-wire shaft. This final reduction employs antibacklash gearing to reduce the indicator deadband to a negligible amount. The slide wire shaft drives the contact arm of the measuring slide wire, the main calibrated scale, and the gearing for coupling to the synchro transmitter. One phase of the motor is energized from the line through a capacitor and the other phase is energized by the amplifier.

Potentiometric System

The major part of the air-borne temperature indicator problem was the development of a good substitute for the dry battery and standard cell. The *5651* voltage reference tube seemed to offer good possibilities of providing this substitute. An exhaustive experimental investigation of its characteristics indicated that:

1. Its voltage regulation characteristic remained quite constant for long periods of time.
2. It had no discontinuities of regulation as did the voltage regulator tubes.
3. Exposure to extreme temperatures did not damage the tube nor permanently change its voltage-current characteristic.
4. While its voltage-current characteristic was displaced with temperature, the shape of the curve remained practically constant, as shown in Fig. 6.
5. The shape of the curve was found to be quite reproducible from tube to tube.

These results led to the conclusion that the *5651* tube could be used in much the same manner as a standard cell if proper temperature compensation were provided or if the *5651* tube were always operated at a constant temperature independent of outside temperature variation.

Since the temperature coefficient of voltage of the *5651* is small, approximately −0.0033 per cent per C, and since it is fairly constant over a considerable range of temperature (Fig. 7), compensation can be accomplished by the use of a series resistor of copper which has a linear positive coefficient. The resultant calculated voltage variation is very small, and experimental results check the calculated curve. Since the resistance of the copper resistor is small, it does not appreciably affect the performance of the *5651* tube.

To ensure close "tracking" as the temperature varies, the *5651* tube must be located close to its compensating resistor and must be thermally "lagged" to match the slower response of the resistor. For reasons discussed later, the potentiometer circuit compartment is thermostatically controlled so that its operating temperature never gets below −30 C. Therefore, close compensation of the *5651* tube below this temperature is not important.

Operated under these conditions, the *5651* tube with its compensating resistor became the basic element of the potentiometric system, of which there are two main parts:

1. A well-regulated d-c power supply, shown schematically in Fig. 8.
2. The potentiometer circuit, shown schematically in Fig. 9.

REGULATED D-C POWER SUPPLY

The direct current for this supply is obtained from the same rectified supply as the direct current used for the balancing system amplifier. Fundamentally, in order to obtain a regulated d-c supply, all that is necessary is to maintain the adjustment of a rheostat placed in series with the unregulated d-c supply so that the voltage across the load remains constant independent of line voltage variations.[2] The method used to accomplish this is shown in Fig. 8. It employs a single tube (*12AX7*) differential-type d-c control amplifier for comparing a direct voltage from the voltage reference tube *5651* with a direct voltage which varies when the a-c line voltage varies. The output

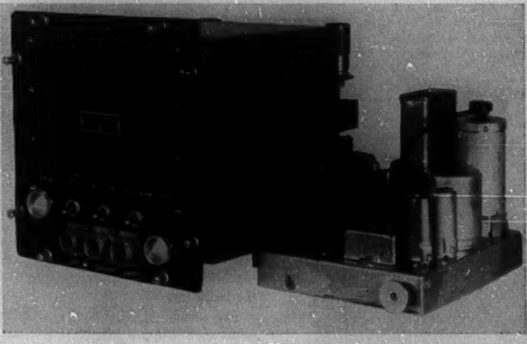


Fig. 4. Indicator out of case, amplifier withdrawn

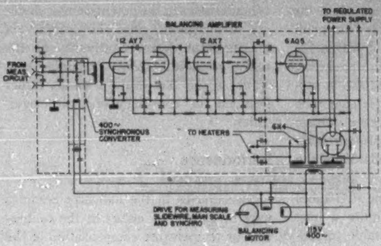

BALANCING AMPLIFIER
TO REGULATED POWER SUPPLY
12 AY7
12 AX7
6AQ5
FROM MEAS CIRCUIT
400~ SYNCHRONOUS CONVERTER
TO HEATERS
6X4
DRIVE FOR MEASURING SLIDEWIRE, MAIN SCALE AND SYNCHRO
BALANCING MOTOR
115V 400~

Fig. 5 (left). The balancing system

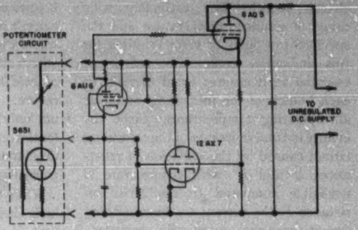

POTENTIOMETER CIRCUIT
6AQ5
6AU6
5651
12AX7
TO UNREGULATED D.C. SUPPLY

Fig. 8 (right). Regulated power supply

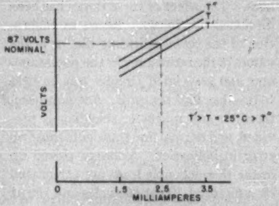


Fig. 6. 5651 voltage-current characteristics at various temperatures

of this stage is amplified by a pentode tube *6AU6* and this amplified output is used to control the resistance of the rheostat tube *6AQ5*.

The output from the rheostat tube, Fig. 8, supplies the direct voltage which is compared with a portion of the reference tube voltage by the *12AX7* tube. It also supplies the current for the parallel combination of the voltage reference tube just mentioned and the load, through a common adjustable series resistance. This series resistance is considerably higher in value than the equivalent resistance of the voltage reference tube-load combination, thus permitting the voltage reference tube to act as a voltage regulator tube and provide a second stage of regulation.

In this application the load remains constant at 10 milliamperes and the voltage reference tube current is 2.5 milliamperes. A change in line voltage of ±10 per cent produces a change in voltage across the *5651* tube of less than ±0.05 per cent. The temperature compensation provided is sufficient to limit the error resulting from changes in ambient temperature of −55 to +55 C to ±0.04 per cent.

The d-c supply used for the balancing system amplifier is connected to the case which is normally grounded. For simplicity, this supply is also used as the source of the potentiometer current. This limits the system to use with ungrounded thermocouples.

POTENTIOMETER CIRCUIT

The potentiometer circuit is shown in Fig. 9. Some of the functions of some parts of the circuit have already been described, i.e., the series resistor (STD ADJ) and the voltage reference tube (*5651* NORM) with its compensating resistor (COMP COIL). The circuit referred to previously as the load consists of an adjustable resistance *F* in series with a 2-branch measuring circuit. A reference junction compensating coil *D* in one branch of the measuring circuit permits a direct indication of temperature.[3] This type of compensation has been widely used in numerous industrial instruments for normal ranges of ambient temperatures. The temperature of the potentiometer circuit compartment of the indicator is thermostatically controlled (mentioned previously) primarily to reduce the range of temperature over which compensation is required. Except for two minor features this measuring circuit is identical to those used in the potentiometer-type Speedomax temperature recorders commonly used in industry. The exceptions are:

1. Use of a continuous-type slide wire in place of the more common wound-type slide wire.
2. The addition, to the slide-wire branch of the circuit, of a coil *D'* which has the same resistance and temperature coefficient as the reference junction coil *D*.

The continuous slide wire *S* is used to achieve better damping and more usable sensitivity than can be obtained with the wound-type slide wire.

The *D'* coil maintains the equality of both branches of the potentiometer circuit. This precaution is unnecessary in industrial instruments, where, under normal operating conditions, the change in the resistance of *D* is negligible compared to the total resistance in each measuring

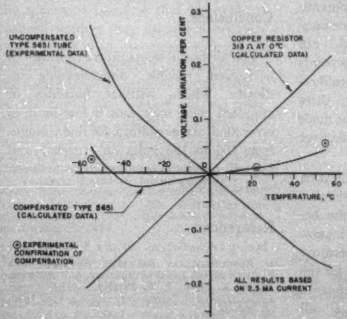


Fig. 7 (left). Temperature compensation of a 5651 tube

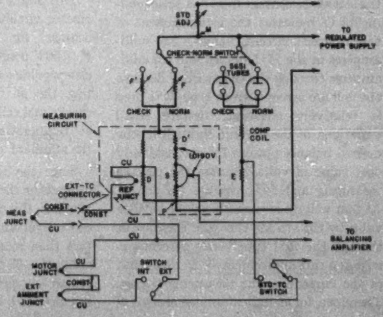

STD ADJ
M
TO REGULATED POWER SUPPLY
CHECK-NORM SWITCH
F'
F
5651 TUBES
CHECK
NORM
CHECK
NORM
COMP COIL
MEASURING CIRCUIT
D'
10190V
CU
D
S
E
EXT-TC CONNECTOR
REF JUNCT
CONST
CONST
MEAS JUNCT
CU
P
CU
CU
TO BALANCING AMPLIFIER
CU
MOTOR JUNCT
CONST
EXT AMBIENT JUNCT
CU
SWITCH
INT
EXT
STD-TC SWITCH

Fig. 9 (right). Schematic diagram of potentiometer circuit

circuit branch. The extremely wide range of ambient temperature and the specified low limit of error required for this indicator make this refinement essential as it ensures equal currents in each branch of the measuring circuit at all operating temperatures. Since the change in total resistance of the measuring circuit caused by the variation of resistances D and D' with temperature is negligible compared to the series resistance F, the current to the measuring circuit remains practically constant. The junction of the constantan thermocouple lead and the copper lead of the measuring circuit is insulated from but fastened to the reference junction coil D to maintain these two components at the same temperature, thus providing for good reference junction compensation.

The system described thus far will maintain a constant current through the potentiometer and, if this current is of the correct value, accurate indicator readings can be obtained over the entire range of ambient temperature. A method of obtaining this correct value will be described next.

For a given value of F the total current through the measuring circuit is governed by the voltage at which the *5651* tube operates. This voltage is not the same for every tube, but if a definite operating current is selected for the *5651* tube, then every tube can be calibrated in terms of the value of F required to ensure a total measuring circuit current of 10 milliamperes. With the rheostat set at its correct value for the particular *5651* tube in use, the current through the measuring circuit can be standardized.

In series with the copper resistor in the *5651* branch is a small resistor E of Evanohm (2.652 ohms). When the STD-TC switch on the front panel is held in the STD position, the self-balancing action of the indicator may be used to compare the voltage across E to the voltage drop from the slide-wire contact to P. A scale reading of O indicates the establishment of the desired current values, 2.5 milliamperes in the *5651* branch and 5.0 milliamperes through the slide wire. Because the voltage across a given type-*5651* tube is relatively constant, despite small variations of the current through it, the desired balance obtains for one, and only one, value of voltage across the parallel measuring and *5651* circuits (M to P). As that voltage is varied from the balance value by the STD ADJ rheostat, the current through the *5651* tube varies percentagewise at a rate approximately 25 times that of the current in the measuring circuit. Therefore, for an indicator limit of error of 0.1 per cent of range (5 microvolts), the current need be adjusted only to a point where the indicator reading is within 125 microvolts, or about three scale divisions, of scale 0 during the standardization procedure. Thus the voltage reference tube not only provides the reference voltage for the regulated supply but also furnishes the standard voltage for standardizing the potentiometer circuit.

While tests showed that the *5651* tubes held their calibration for quite a long time, it was considered desirable to provide some convenient method of making a quick check of this standard voltage. For this purpose a second *5651* tube (CHECK) and its associated rheostat F' are supplied. This rheostat is, of course, set to the correct value for the CHECK *5651* tube. If a constant input signal is applied to the indicator, the stability of the calibration of the NORM *5651* tube can be checked by changing the CHECK-NORM switch from the NORM position to the CHECK position. If, after the circuit has been standardized by means of the STD-TC switch and the STD ADJ rheostat, the indicator reading of the input signal is the same for the CHECK position as for the NORM position, the indicator is correctly calibrated. A difference in the two readings indicates that the characteristics of at least one of the *5651* tubes have changed and both tubes should be replaced. While it is possible to obtain the same indicator reading when both tubes are in error, this condition is not probable.

A relatively constant input signal for the preceding test is obtained from two thermocouples connected in series opposition, with one couple in close thermal contact with the hot driving motor and the second couple exposed to the colder ambient temperature. This signal is connected to the input of the indicator by putting the INT-TC-EXT switch in the INT position.

Tap points are provided in the potentiometer circuit for checking its accuracy against an external standard cell (1.0190 V points of Fig. 9). These points also provide a means of accurately calibrating the *5651* tubes. However, these checks and calibrations can only be made at temperatures that will not damage the standard cell.

Since the accuracy of standardization depends on the constancy of the resistance of coil E, this coil is made of Evanohm because Evanohm varies less with temperature than manganin over a temperature range of −55 to +55 C.[4] Except for the copper compensating coil, the reference junction coil, and its counterpart D', all other resistors in the potentiometer circuit are of manganin, since the accuracy of measurement depends primarily on the ratios of these resistors remaining constant as the temperature varies rather than on the constancy of their magnitudes.

Performance

Inaccuracies in the indicator readings may be caused by errors in the calibration or temperature compensation of the *5651* tubes and by variations in line voltages. The effect of these errors has been discussed previously. Errors in calibration of the main indicating dial or in the values of the resistances in the potentiometer and associated circuits can be held to less than 0.05 per cent. Adequate electrostatic and magnetic shielding is provided, and normal line transients cause no error in indication. Voltages picked up in the thermocouple leads are attenuated by a factor of 100 at 60 cycles and 700 at 400 cycles.

Performance tests over the entire range of ambient temperatures (−55 to +55 C), line voltage (105 to 125 volts), and line frequency (380 to 420 cycles) show that the instrument

1. Operating as a millivolt indicator (without reference junction compensation) has a resultant error of only ±0.1 per cent.
2. Operating as a temperature indicator (with reference junction compensation and a measuring range of −70 to +60 C) has a resultant error of less than ±0.9 C.

Full-scale balancing time at ambient temperatures from approximately 0 to +55 C is less than 4 seconds. At temperatures below 0 C the balancing time increases as the temperature decreases, but never becomes slower than 10 seconds. The indicator weighs 24 pounds and will withstand the severe shock and vibration tests specified by MIL-E-5272-A.[5]

Conclusions

The air-borne temperature indicator is a rugged, reliable, direct-reading indicator, which, without batteries or standard cells, maintains a low limit of error at all scale readings, for line voltages from 105 to 125 volts (400 cycles) and for all ambient temperatures between −55 and +55 C.

References

1. ELECTRONICALLY BALANCED RECORDER FOR FLIGHT TESTING AND SPECTROSCOPY, A. J. Williams, Jr., W. R. Clark, R. E. Tarpley. *AIEE Transactions (Electrical Engineering)*, vol. 65, April 1946, pp. 205–08.

Vo
Pt.
19

2. Stabilized Power Supplies for Instrument Applications, W. G. Amey, F. H. Krantz, W. R. Clark, A. J. Williams, Jr. *AIEE Transactions*, vol. 73, pt. 1, Nov. 1954, pp. 484–89.

3. Pyrometric Practice, P. D. Foote, C. O. Fairchild, T. R. Harrison. *Technological Paper No. 170*, National Bureau of Standards, Washington, D. C., Feb. 16, 1921, pp. 58–59.

4. Conference on Electrically Operated Recording and Controlling Instruments. *AIEE Special Publication S-52*, April 1953, "Reliability of Components Used in Recording and Controlling Instruments," W. R. Clark, pp. 37–52.

5. General Specification for Environmental Testing, Aeronautical and Associated Equipment. *MIL-E-527A*, U. S. Air Force, Washington, D. C., Sept. 16, 1952.

On the Optimum Design of Cathode Followers

L. M. VALLESE
ASSOCIATE MEMBER AIEE

CATHODE followers find wide application in modern electronics because of their simple structure, their fairly distortionless operation, and their low output impedance. Designing for small signal amplification is easy; for large signal amplification a complication is introduced with the requirement that the distortion-free driving range of the control grid be utilized fully. This problem can be solved graphically, but the procedure is cumbersome and of limited application. Its usefulness can be greatly enhanced if it is used merely to check and improve an approximate preliminary design obtained analytically.

For the derivation of such a design the family of plate characteristics of the tube, which is usually a triode, can be represented by means of the bilinear relationship

$$i_b = g_p(e_b + \mu e_{ck} - V_0') \qquad (1)$$

This equation corresponds to a set of straight lines (Fig. 1), having slope g_p and intercept $V_{bk} = -\mu e_{ck} + V_0'$ with the $i_b = 0$ axis; in particular V_0' is the intercept corresponding to $e_{ck} = 0$.

The main purpose of the design consists in determining the values of E_Q, I_Q at the quiescent point so that the grid-to-cathode voltage may be varied between -0.5 V and a value corresponding to a minimum plate current I_{bm}. In addition, the design may be optimized with respect to the power and voltage output, plate dissipation, etc. In the following the cases of a-c and d-c amplifiers will be discussed separately; it will be shown that the design of a-c amplifiers is subject to less restriction than that of d-c amplifiers. The circuit element in the cathode lead of a-c amplifiers can have either a conductive admittance $G_k + j0$ or an inductive admittance $0 - jB_k$, while the corresponding element of d-c amplifiers must have conductive admittance.

A-C Cathode Followers with Conductive Cathode Lead Element

The circuit diagram of a typical a-c cathode follower with conductive cathode lead element is indicated in Fig. 2. The analysis will be limited to the mid-frequency range, disregarding high- and low-frequency effects. For this reason the interelectrode and coupling capacitances will be neglected and the design will be limited to the determination of the quantities G_k, G_L, E_{bb}, E_0, E_s, $(E_{ck})_Q$, E_Q, I_Q. The interpretation of these symbols follows from Figs. 1 and 2; it should be noticed that E_0 and E_s are peak rather than rms values.

The following limitations must be satisfied

$$\left.\begin{aligned} &e_{ck} \le -0.5V \\ &I_{bm} \le i_b \le I_{k\,\max} \\ &E_{bb} \le E_{bb\,\max} \\ &E_Q I_Q \le P_{p\,\max} \end{aligned}\right\} \qquad (2)$$

The last equation applies for zero signal dissipation. Indicating with I_{bM}, the maximum value of the plate current ($I_{bM} \le I_{k\,\max}$) there follows

$$2I_Q = I_{bm} + I_{bM} \qquad (3)$$

$$I_{bM} = g_p(E_Q - E_0 - V_{bk}') \qquad (4)$$

Equation 4 represents a point of the characteristic of i_b versus e_b for $E_{ck} = -0.5$ V and is obtained from equation 1 letting $e_b = E_Q - E_0$ and $V_{bk} = V_{bk}'$. Furthermore

$$I_Q = G_k(E_{bb} - E_Q) = I_{bm} + E_0(G_k + G_L) \qquad (5)$$

From these equations there follows

$$E_Q = V_{bk}' + I_{bm}/g_p + E_0$$

$$E_{bb} = E_Q + I_Q/G_k$$

Finally from the equation $V_{bk} = -\mu e_{ck} + V_0'$ there follows

$$(\mu E_{ck})_Q = V_0' - E_Q + I_Q/g_p$$

The peak input voltage is

$$E_s = E_0 + (E_{ck})_Q - 0.5$$

By elimination of E_Q, I_Q, I_{bm} from equations 3, 4, and 5 one has

$$G_k^2 - \frac{g_p}{2E_0}\left(E_{bb} - 2E_0\left(1 + \frac{G_L}{g_p}\right) - \frac{I_{bM}}{g_p} - V_{bk}'\right)G_k + \frac{g_p}{2}\left(G_L + \frac{I_{bm}}{E_0}\right) = 0$$

which is a relationship among E_{bb}, E_0, G_k, G_L and can be used to determine one

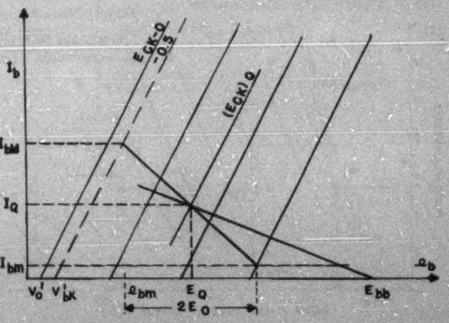


Fig. 1. Linearized plate characteristics of a triode vacuum tube

---

Paper 55-7, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted August 17, 1954; made available for printing October 27, 1954.

L. M. Vallese is with the Polytechnic Institute of Brooklyn, Brooklyn, N. Y.

This paper is an abstract of report *R-391-54* prepared at the Microwave Research Institute of the Polytechnic Institute of Brooklyn under a contract sponsored by the Office of Naval Research.

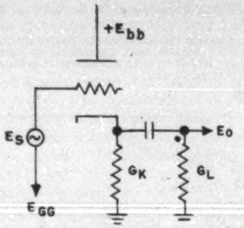


Fig. 2. A-c cathode follower with conductive cathode admittance

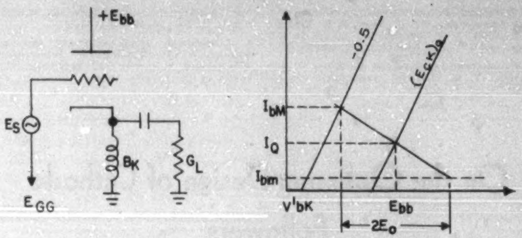


Fig. 4. A-c cathode follower with susceptive cathode admittance

Vol.
Pt. 1
1955

of these quantities when the other three are given.

$$G_L=\frac{g_p(E_{bb}-V_{bk}'-2E_0)G_k-2E_0G_k^2-I_{bm}(g_p+G_k)}{E_0(2G_k+g_p)}$$

or

$$E_0=\frac{g_p(E_{bb}-V_{bk}')G_k-(g_p+G_k)I_{bm}}{2G_k(G_k+g_p)+(g_p+2G_k)G_L} \qquad (6)$$

or

$$G_k=\frac{g_p}{4E_0}\left[E_{bb}-2E_0\left(1+\frac{G_L}{g_p}\right)-\frac{I_{bm}}{g_p}-V_{bk}'\right]\pm \sqrt{\frac{g_p^2}{16E_0^2}\left[E_{bb}-2E_0\left(1+\frac{G_L}{g_p}\right)-\frac{I_{bm}}{g_p}-V_{bk}'\right]^2-\frac{g_p}{2}\left(G_L+\frac{I_{bm}}{E_0}\right)} \qquad (7)$$

or finally

$$E_{bb}=V_{bk}'+I_{bm}\left(\frac{1}{G_k}+\frac{1}{g_p}\right)+2E_0\left(1+\frac{G_L}{2G_k}+\frac{G_L+G_k}{g_p}\right) \qquad (8)$$

Two other important relations are

$$I_{bM}=I_{bm}+2E_0(G_k+G_L)=\frac{2(G_k+G_L)g_pG_k(E_{bb}-V_{bk}')-g_pG_LI_{bm}}{2(g_p+G_k)(G_k+G_L)-g_pG_L} \qquad (9)$$

and

$$P_p=E_QI_Q=\left[E_0\left(1+\frac{2(G_k+G_L)}{g_k}\right)+V_{bk}'+\frac{I_{bm}}{g_p}\right][I_{bm}+E_0(G_k+G_L)]=\left[E_{bb}-\frac{1}{G_k}(I_{bm}+E_0(G_k+G_L)\right][I_{bm}+E_0(G_k+G_L)] \qquad (10)$$

The previous equations will now be applied to a number of practical problems. In particular, numerical design values will be computed for circuits using the *6J5*. From the linearized plate characteristics of this tube (Fig. 3)

$g_p=0.13\times10^{-3}$ mhos, $V_0'=30\ V$, $V_{bk}'=40\ V$
$I_{bm}=0.5\ mA$, $\mu=20$, $I_{k\ \max}=20\ mA$, $E_{bb\ \max}=300\ V$
$P_{p\ \max}=2.5\ w$

DESIGN NO. 1

Given the tube G_k, G_L, E_0, find the minimum necessary value of E_{bb}. Equation 8 is used. That the values of E_{bb}, I_{bm}, and P_p are less than the limit values (equation 2) is checked. Finally $(E_{ck})_Q$ E_{GG}, and E_s are computed.

For example, for a *6J5* tube with $G_k=0.4$ times 10^{-3} mhos, $G_L=10^{-3}$ mhos, $E_0=5\ V$, there follows $E_{bb}=173.5\ V$; in addition, $I_{bM}=14.5\ mA$, $I_Q=7.5\ mA$, $E_Q=155\ V$, $P_p=1.17\ w$, $(E_{ck})_Q=-3.3\ V$, $E_{GG}=151.7\ V$, $E_s=7.7$ V.

DESIGN NO. 2

Given the tube G_k, G_L, find E_{bb} for max E_0 permissible. Since E_0 increases linearly with E_{bb}, it is sufficient to investigate the maximum value of E_0. With respect to the current limitation in equation 2, this is (equation 9)

$$E_0'=\frac{I_{kM}-I_{bm}}{2(G_k+G_L)} \qquad (11)$$

With respect to the power limitation (equation 2), there is from equation 10

$$E_0''=-a+\sqrt{a^2+b} \qquad (12)$$

where

$$a=\frac{[g_p+2(G_k+G_L)]I_{bm}+(G_k+G_L)(g_pV_{bk}'+I_{bm})}{2(G_k+G_L)[g_p+2(G_k+G_L)]}$$

$$b=\frac{g_pP_{pM}-(g_pV_{bk}'+I_{bm})I_{bm}}{(G_k+G_L)[g_p+2(G_k+G_L)]}$$

The lower of the values, equations 11 and 12 is the maximum permissible E_0. Correspondingly, E_{bb} is obtained from equation 8.

For example, for a *6J5* tube with $G_k=0.4$ times 10^{-3} mhos, $G_L=10^{-3}$ mhos, there results $E_0'=7.3\ V$, $E_0''=7.8\ V$.

DESIGN NO. 3

Given the tube G_L, E_0, E_{bb}, find G_k. Equation 7 is used, providing two solutions for G_k. These are checked with respect to the limitations (equation 2). Assuming that both G_k values are permissible, the final choice will be determined on the basis of such considerations as plate power dissipation, peak cathode current, and output admittance.

For example, for a *6J5* tube, given $G_L=10^{-3}$ mhos, $E_0=5\ V$, $E_{bb}=173\ V$, the corresponding values for G_k are as follows

$G_k'=0.4\times10^{-3}$ mhos

$G_k''=0.177\times10^{-3}$ mhos

Both these values are permissible.

DESIGN NO. 4

Given the tube G_L, E_{bb}, find G_k for max E_0. From equation 6, letting $dE_0/dG_k=0$

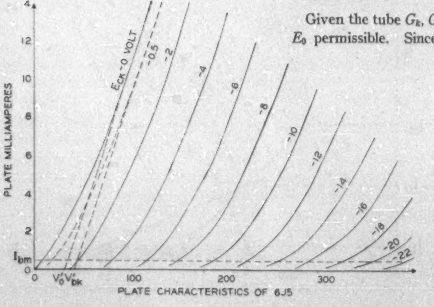


Fig. 3. Plate characteristics of 6J5 tube

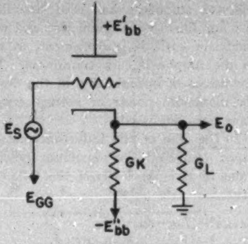


Fig. 5. D-c cathode follower

$$G_k=\frac{g_pI_{bm}}{g_p(E_{bb}-V_{bk}')-I_{bm}}+\sqrt{\frac{g_p^2I_{bm}^2}{(g_p(E_{bb}-V_{bk}')-I_{bm})^2}+\frac{g_p}{2}\left(G_L+\frac{2(g_p+G_L)I_{bm}}{g_p(E_{bb}-V_{bk}')-I_{bm}}\right)}$$

For example, letting $G_L=10^{-3}$ mhos, $E_{bb}=173$ *V*, there results for the *6J5* tube $G_k=2.6$ times 10^{-4} mhos and $E_0=5.09$ *V*.

DESIGN No. 5

Given the tube and E_{bb}, find G_k and G_L for maximum output power $P_L=G_LE_0^2/2$.

This design may be considered of academic interest, since power amplifiers are more efficient with an inductive cathode lead element. However, if the high-frequency response is important, the question of the efficiency might become secondary.

According to equation 6

$$P_L=\frac{G_L}{2}\left[\frac{g_p(E_{bb}-V_{bk}')G_k-(g_p+G_k)I_{bm}}{2G_k(G_k+g_p)+(g_p+2G_k)G_L}\right]^2$$

which is a function of E_{bb}, G_k and G_L. If either G_k or G_L are given, the optimum value of the conductance G_L or G_k is determined respectively with the condition

$$\frac{dP_L}{dG_L}=0 \quad \text{or} \quad \frac{dP_L}{dG_k}=0 \tag{13}$$

If G_k and G_L are both undetermined, the latter two equations must be solved simultaneously. Letting $I_{bm}=0$ to simplify the final expressions, there follows, for G_k given

$$G_{L\ \text{opt}}=2G_k(G_k+g_p)/(2G_k+g_p)$$

and for G_L given

$$G_{k\text{opt}}=\sqrt{g_pG_L/2}$$

Finally, if both G_k and G_L are undetermined, their optimum values are

$$G_k=g_p/\sqrt{2} \qquad G_L=g_p \tag{14}$$

Correspondingly, the peak power output is

$$P_{L\ \text{peak}}=\frac{g_p(E_{bb}-V_{bk}')^2}{16(1+\sqrt{2})^2}$$

It is of interest to point out that the result (equation 14) applies also for grounded cathode amplifiers.

A-C Cathode Followers with Inductive Cathode Lead Element

The circuit diagram of Fig. 4 is now studied; it is assumed that $jB_k=1/j\omega L$ is negligible. The limitations of the design are still expressed by equations 2, where $E_Q=E_{bb}$. The design in the present case is much simplified and reduces to finding I_Q, since E_Q is fixed. The design equations are

$$I_Q=I_{bm}+G_LE_0$$

$$I_{bM}=g_p(E_{bb}-V_{bk}'-E_0)=2I_Q-I_{bm}$$

$$\mu(E_{ck})_Q=(V_0'-E_{bb}+I_Q/g_p)$$

$$E_s=E_0+(E_{ck})_Q-0.5$$

Eliminating I_Q between the first and the second equation results in

$$g_p(E_{bb}-V_{bk}')-I_{bm}=E_0(2G_L+g_p) \tag{15}$$

From this equation, if E_{bb} and G_L are given

$$E_0=\frac{g_p(E_{bb}-V_{bk}')-I_{bm}}{2G_L+g_p} \tag{16}$$

and if E_{bb} and E_0 are given

$$G_L=\frac{g_p(E_{bb}-V_{bk}'-E_0)-I_{bm}}{2E_0}$$

The maximum possible value of E_0 for given E_{bb} occurs when $G_L=0$ and is

$$E_{0\ \max}=E_{bb}-V_{bk}'-I_{bm}/g_p$$

On the other hand, when $G_L\neq0$, the maximum possible value of E_0 is

$$E_{0\ \max}=(I_{kM}-I_{bm})/2G_L$$

according to equations 15 and 16.

As an example, given a *6J5* tube with $E_{bb}=200$ *V*, $G_L=10^{-3}$ mhos, there follows $E_0=9.6$ *V*, $I_Q=10.1$ *mA*, $(E_{ck})_Q=-9.7$ *V*. The maximum possible value of E_0 for the given value of G_L is 9.75 *V*; the maximum possible value of E_0 for $G_L=0$ is 156 *V*.

An interesting problem is the determination of the optimum value of G_L for maximum power output.

$$P_L=G_LE_0^2/2=G_L(g_p(E_{bb}-V_{bk}')-I_{bm})^2/2(2G_L+g_p)^2$$

and letting $dP_L/dG_L=0$, there follows $2G_L=g_p$, which is the same result that applies for grounded cathode amplifiers. Correspondingly, the maximum power output is

$$P_{L\ \text{peak}}=\frac{g_p(E_{bb}-V_{bk}')^2}{16}$$

D-C Cathode Followers

A typical circuit for the d-c cathode follower is indicated in Fig. 5. For generality it has been assumed that $E_{bb}'\neq E_{bb}''$.

One of the conditions of the design is that the cathode be at ground potential. Correspondingly

$$E_Q=E_{bb}' \qquad I_Q=G_kE_{bb}'' \tag{17}$$

By application of Thévenin's theorem the circuit of Fig. 5 can be modified into that of Fig. 6(A), with

$$G'=G_k+G_L \qquad E'=G_kE_{bb}''/(G_k+G_L)$$

The latter circuit is equivalent to that of Fig. 6(B) which is similar to the circuit of an a-c cathode follower with zero load admittance G_L.

The limit conditions of operation are

$$E_{ck}\leq-0.5\ V$$

$$I_Q=G_kE_{bb}''\leq(I_{kM}+I_{bm})/2 \tag{18}$$

$$P_p=E_{bb}'I_Q\leq P_{p\ \max} \tag{19}$$

The design equations are

$$E_{bb}'=E_0+V_{bk}'+I_{bM}/g_p$$

$$=V_{bk}'+E_0(1+2(G_k+G_L)/g_p) \tag{20}$$

$$I_Q=[I_{bm}+g_p(E_{bb}'-V_{bk}'-E_0)]/2 \tag{21}$$

$$=I_{bm}+E_0(G_k+G_L) \tag{22}$$

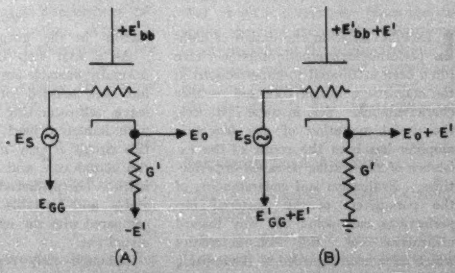


Fig. 6. D-c cathode follower

$$\mu(E_{ck})_Q = V_0' - E_{bb}' + I_Q/g_p \qquad (23)$$

$$E_s = E_0 + (E_{ck})_Q - 0.5 \qquad (24)$$

In general E_0 and G_L are prescribed, and the purpose of the design consists in choosing tube and evaluating E_{bb}', E_{bb}'', and G_k.

The choice of the tube is determined by the maximum voltage output E_0 that it can deliver in accordance with the limitations (equations 18 and 19). The upper limit of E_0 is obtained letting G_k approach zero. Combining equations 18 and 22

$$E_{0\,\max}' = (I_{kM} - I_{bm})/2G_L \qquad (25)$$

On the other hand, combining equations 19, 20, and 22, there results

$$[E_0(1+2(G_k+G_L)/g_p) + V_{bk}'](E_0(G_k+G_L) + I_{bm}) \le P_{pM}$$

wherefrom there follows

$$E_{0\max}'' = \frac{-g_p V_{bk}'}{2(g_p+2G_L)} + \sqrt{\frac{g_p P_{p\max}}{G_L(g_p+2G_L)} + \left(\frac{g_p V_{bk}'}{2(g_p+2G_L)}\right)^2} \qquad (26)$$

The lower of the values equations 25 and 26 should be larger than the required E_0.

After the tube has been selected, E_{bb}' is evaluated. The upper limit of E_{bb}' again is determined in accordance with the limitations (equations 18 and 19). Combining equations 18 and 21

$$E_{bb}' \le E_0 + V_{bk}' + (I_{kM} - I_{bm})/g_p \qquad (27)$$

On the other hand, combining equations 19 and 21, there results

$$E_{bb}' \le \frac{1}{2}\left(V_{bk}' + E_0 - \frac{2I_{bm}}{g_p}\right) + \sqrt{\frac{1}{4}\left(V_{bk}' + E_0 - \frac{2I_{bm}}{g_p}\right)^2 + \frac{2P_{pM}}{g_p}} \qquad (28)$$

In general, the value of E_{bb}' is chosen in dependence of the available power supply, so that the limitations (equations 27 and 28) are satisfied.

After the value of E_{bb}' has been determined, equations 21 and 22 provide respectively I_Q and G_k. Finally, the value of E_{bb}'' is obtained from equation 17, and the values of $(E_{ck})_Q$ and E_s are computed from equation 23 and 24 respectively.

For example, consider the design of a d-c cathode follower with $E_0 = 8$ V, $G_L = 10^{-3}$ mhos. For the tube *6J5* equations 25 and 26 provide respectively

$$E_{0\max}' = 9.75 \text{ V}$$

$$E_{0\max}'' = 11.2 \text{ V}$$

Therefore the *6J5* tube is suited for the present design. According to equations 27 and 28, the upper limit of E_{bb}' is the lower of the following values: $E_{bb}' \le 196$ V, $E_{bb}' \le 214$ V.

Letting $E_{bb}' = 180$ V, equation 21 provides $I_Q = 9$ mA, and equation 22 gives $G_k = 0.06$ times 10^{-3} mhos. Finally $E_{bb}'' = 150$ V and $(E_{ck})_Q = -4$ V, and $E_s = 11.5$ V.

Summary

An analytical procedure to obtain a preliminary design of large signal cathode follower amplifiers has been described. A-c amplifiers with conductive and with inductive cathode lead element, as well as d-c amplifiers, are considered. The conditions of optimization of the design for maximum power or voltage output have been discussed.

In the case of the inductance-capacitance amplifier the maximum possible value of E_0 is the lower of the following ones

$$E_{0\max}' = (I_{kM} - I_{bm})/2G_L$$

$$E_{0\max}'' = (P_{p\max} - E_{bb}I_{bm})/G_L E_{bb}$$

In the case of the d-c amplifier the design for max P_L gives

$$G_k \cong 0,\ G_L = G_k + g_p/2$$

$$P_{L\max} = g_p(E_{bb}' - V_{bk}')^2/_{16}$$

References

1. The Cathode Follower Stage, E. Williams. *Wireless World and Radio Review*, London, England, July 1941, p. 175.

2. Cathode Follower Circuits, K. Schlesinger. *Proceedings*, Institute of Radio Engineers, New York, N. Y. Dec. 1945, p. 843.

3. A Graphical Analysis of the Cathode-Coupled Amplifier, M. S. Rifkin. *Communications*, New York, N. Y., Dec. 1946, p. 16.

4. Valves with Resistive Loads, S. W. Amos. *Westinghouse Engineer*, Pittsburgh, Pa., April 1949, p. 119.

5. Vacuum Tube Amplifiers (book), edited by George E. Valley, Jr., Henry Wallman. McGraw-Hill Book Company, New York, N. Y., 1948.

6. A Practical Cathode Follower Audio Amplifier, W. F. Gibson, R. Pavlat. *Audio Engineering*, Mineola, N. Y., May 1949, p. 9.

7. The Theory and Design of Cathode Follower Output Stages, E. T. Emms. *Electronic Engineering*, London, England, Sept. 1953.

Inductive Load Instability in Magnetic Amplifiers

H. I. LEON
NONMEMBER AIEE

ALLEN B. ROSENSTIEN
ASSOCIATE MEMBER AIEE

LIMITATIONS in the design of high-performance magnetic amplifiers have often been attributed to imperfections in the magnetic amplifier core and rectifier characteristics. Yet, in some instances, the useful operation of the magnetic amplifer has been the result of the existence of these same so-called imperfections. Evaluation and understanding of the parameters of the material imperfections can point the way toward utilization of the characteristics which have been regarded as the limiting factors in the amplifier performance.

As a first step in this direction, a normally unstable amplifier configuration has been selected for analysis, the full-wave, self-saturating magnetic amplifier with inductive load. The behavior of this circuit is very much dependent on the actual core and rectifier characteristics. The predicted boundary between stable and unstable operation will be compared with the experimentally determined one.

Although early reports on magnetic amplifiers have noted instabilities in circuits with ostensibly less than 100-percent feedback, it is only recently that articles have begun to appear which describe the mechanism of these instabilities. The effects of core material on stability has been investigated.[1] Instability caused by inductive loads has been mentioned in several articles[2-6] and has been avoided in one of several ways. The inductive load is tuned to resonance to present a capacitive reactance; a rectifier or rectifier bridge is added to the load

---

Paper 55-21, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted September 27, 1954; made available for printing November 10, 1954.

H. I. Leon and Allen B. Rosenstien are with the University of California, Los Angeles, Calif.

The research described in this paper was sponsored by the Aviation Ordnance Department, U. S. Naval Ordnance Test Station, Inyokern, Calif.

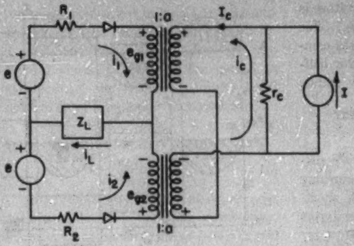


Fig. 1 (left). Schematic diagram of magnetic amplifier circuit showing sign convention

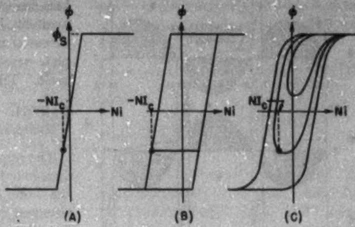


Fig. 2 (right). Postulated magnetization characteristics of magnetic cores

l, 74
. 1
55

circuit; the amplifier circuit is changed to a more stable configuration; or the control circuit constraint is made very high or inductive loads are avoided altogether.

The purpose of this paper is to examine the mechanism of the inductive instability to predict the unstable domain and to seek more efficient stability methods. To develop the practical magnetic amplifier circuit analysis, it has been expedient to analyze first the completely idealized amplifier utilizing "perfect" rectifiers and magnetic cores. Then one idealized circuit element or material characteristic at a time has been changed to a more realistic one and the system reanalyzed. This procedure has been continued until practically all of the nonlinearities of the amplifier elements have been considered and the performance predicted from the mathematical model of the system offered a reasonable representation of the actual system performance.

Nomenclature

C = capacitance
E = direct voltage, peak alternating voltage
H = magnetizing force
I = direct current
L = inductance
N = number of turns
R = resistance from gate side
R_b = loop back resistance
R_e = reflected control circuit resistance
R_f = loop forward resistance
T = time constant
$Q = L/R$
Z = impedance
a = turns ratio
e = voltage
i = current
i_L = load current
n = integer
r = resistance
t = time
x = angle
α = ratio
β = feedback ratio
γ = ratio
ϕ = flux
θ = angle
ω = angular frequency
τ = time constant
$\Delta(\ )$ = change in $(\ )$

Subscripts

ac = alternating current
c = control, coercive
g = gate
s = saturation
1 = loop 1
2 = loop 2

Idealized Circuit Analyses

Consider the operation of the idealized full-wave, self-saturating magnetic amplifier, as shown in Fig. 1. The rectifiers have zero forward resistance, infinite back resistance, and no capacitance. The magnetic cores have either the single-valued characteristic, shown in Fig. 2(A), or the double-valued characteristic, shown in Fig. 2(B). The slopes of these characteristics are either zero or slightly positive in the saturated zones, and large (but finite) in the unsaturated zones. Loop resistances R_1 and R_2 are negligible.

The core winding in the load circuit is called the gate winding, the other winding is called the control winding. The term gate may refer to either the core or the winding. Whenever a voltage appears across the gate winding, the core flux must change in accordance with

$$e_g = N_g \frac{d\phi}{dt} \tag{1}$$

The change in flux is

$$\Delta\phi = \phi(t_2) - \phi(t_1) = \frac{1}{N}\int_{t_1}^{t_2} e\,dt \tag{2}$$

A change of flux may be referred to as a $\Delta\phi$ or its equivalent volt-seconds as calculated from this equation. A core is said to absorb voltage whenever the core flux is changing.

For a sinusoidal voltage, the change in flux for a half-cycle is

$$\Delta\phi = \int_0^{\pi/\omega} E \sin \omega t\,dt = \frac{2E}{N\omega} \tag{3}$$

If the change in flux is from full negative to full positive saturation, then

$$E = N\omega\phi_s \tag{4}$$

where ϕ_s is the saturation flux. In this analysis, the supply voltage will be adjusted to this value.

Whenever a gate saturates, it will be considered as a short circuit across any of its windings. Whenever a gate is not saturated, it will be considered as a perfect transformer. When a core is changing flux in the negative direction, i.e., so that negative loop current will attempt to flow if the core reaches negative saturation, it will be called the reset interval for that core. When a core is changing flux in the positive direction, it will be called the set interval for that core. A core is completely reset when the flux has reached the bias value, i.e., the value of flux corresponding to the d-c ampere-turns in the control winding on the descending branch of the hysteresis loop. A completely reset core will have zero current in its gate winding. A gate voltage that tends to reset the core further will cause a current flow in the reverse rectifier direction.

The criterion for stability to be used is that the steady-state load current shall change incrementally for an incremental change in load current. Six magnetic amplifier circuits having progressively fewer ideal elements or characteristics will be analyzed to determine the factors which influence inductive load instability. In the description of the last circuit, case VI, a practical magnetic amplifier circuit having a reactive load and finite control-circuit resistance is discussed qualitatively. Then a quantitative analysis of that circuit is presented. Since the unstable region of operation usually occurs experimentally at the minimum load current point, only small load currents and low firing angles will be considered in all cases in order to find the boundary of stability.

Resistance Load Analyses

Case I. Nonzero Control-Circuit Resistance

For the bias point, shown in Fig. 2 (not quite to $-\phi_s$), wave forms of gate

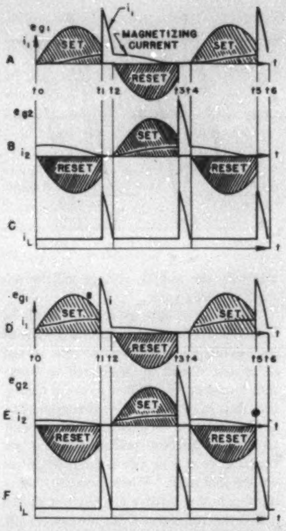


Fig. 3. Wave forms for cases I and II, a resistance load and nonzero control-circuit resistance with (D, E, F) and without (A, B, C) magnetic hysteresis

voltage and load current are shown in Fig. 3. Curves *A*, *B*, and *C* are for the single-valued characteristic of Fig. 2(A) and curves *D*, *E*, and *F* are for the double-valued characteristic of Fig. 2(B). The initial conditions have gate 1 starting at the bias flux and gate 2 at positive saturation.

From t_0 to t_1, gate 1 sets and gate 2 resets. Magnetizing currents flow in both gate windings. At t_1, gate 1 saturates and $i_1=e_1/R_L$. Double reset voltage is now impressed on loop 2 as saturated gate 1 is considered a short circuit. However, gate 2 has reset completely at t_1 and the rectifier blocks any negative current preventing any further change in flux.

At t_2, the current i_1 decreases to the value at which gate 1 emerges from the saturated zone. Gate 1 now is subjected to reset volt-seconds and gate 2 to set volt-seconds. No control-circuit current is induced because the voltages across both gates are always equal and opposite.

The set volt-seconds are equal to the area under the e_{g1} curve from t_0 to t_1. The reset volt-seconds are equal to the area under the e_{g2} curve in the same interval. If, for some reason, reset was inhibited so that gate 2 did not attain the bias flux at t_1, then double reset voltage is available in the interval t_1 to t_2. This excess of reset volt-seconds insures complete reset before the end of the half-cycle. This circuit is inherently stable.

CASE II. ZERO CONTROL-CIRCUIT RESISTANCE

The zero value of control-circuit resistance may be obtained without disturbing the value of i_c by an additional, short-circuited control winding. Wave forms for this case are identical with those for case I. Operation differs in only one respect: during the interval t_1 to t_2, gate 1 is saturated and short-circuits gate 2 from the control side. At t_2, however, gate 2 is completely reset, and the rectifier blocks any reverse current. There is consequently no change in steady-state circuit operation from case I.

Inductive Load Analyses

Cases I and II have shown that, with resistive loading, the set and reset volt-seconds are always equal and the amplifier is inherently stable. The essential elements of the inductive load instability can be easily visualized from Fig. 2. Consider the commutation time interval when the load current is transferring from loop 1 to loop 2. If the load is inductive, a voltage is induced in the inductance that attempts to maintain the current flow at previous magnitudes. If the polarities of the voltages are noted around both loop 1 and loop 2, it can be seen that the voltage across the inductance adds additional set volt-seconds to gate 2 and subtracts from the reset volt-seconds available to gate 1. In contrast, the polarity across a capacitive load at commutation is exactly opposite from that of the inductive load with the result that set volt-seconds are reduced and reset volt-seconds increased. Capacitive loading is, therefore, inherently quite stable.

CASE III. PURE INDUCTANCE LOAD AND ZERO CONTROL-CIRCUIT RESISTANCE

The case of pure inductance load and zero control-circuit resistance cannot be constructed with practical materials. The theoretical wave forms are given in Fig. 4 for completeness. At t_0, gate 1 is started at the bias point and gate 2 at positive saturation. From t_0 to t_1, gate 1 sets and gate 2 resets. Wave forms for negligible magnetizing currents are shown in Fig. 4(A), (B), and (C). At t_1, gate 1 saturates and gate 2 is completely reset. During the interval t_1 to t_2, the rectifier prevents any change in loop 2. Load current flows in loop 1 as

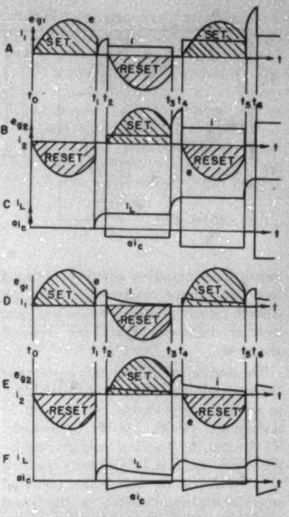


Fig. 4. Wave forms for cases III and IV (A, B, C), inductance load and zero control-circuit resistance, and for case IV (D, E, F), inductance-resistance load and zero control circuit resistance

$$i_L(t)=i_1(t_1)+\frac{E}{\omega L}(\cos\omega t_1-\cos\omega t);\quad t_1<t<t_2 \tag{5}$$

At t_2, the supply voltage changes sign and set voltage is applied to loop 2. At this instant, load current is quite large, in fact at its maximum value, and gate 1 remains saturated. The control winding of gate 2 is short-circuited by the control winding on gate 1. Load current immediately divides between loops 1 and 2; see Appendix III. Control-circuit current flows, equal and opposite to the loop current, so that the net ampere-turns on each gate are zero (except for magnetizing current).

Gate 1 resets and gate 2 sets at normal voltage starting from t_2 independently of the load current which is maintained constant from t_2 to t_3 by the dissipationless, inductive circuit; see Appendix III. At t_3, gate 2 saturates and gate 1 flux has attained the bias value. However, the rectifier in loop 1 cannot block because the rectifier carries a large positive current. The voltage in loop 1 is negative and has double amplitude because gate 2 is a short circuit. Hence, the current in the outside impedanceless loop, i.e., through both gates and power supplies, immediately jumps to the value of the loop current, reducing the current in loop 1 to zero and transferring the total load current to loop 2. The rectifier in loop 1 prevents any further change in loop 1 and gate 1 is clamped at the bias-flux point. Load current from loop 2 now increases as

$$i_L(t)=i_2(t_3)+\frac{E}{\omega L}(\cos\omega t_3-\cos\omega t);\ t_3<t<t_4 \tag{6}$$

At t_4, the supply voltage again changes sign and loop 1 becomes conducting. However, the control side of gate 1 is still short-circuited and, as before, the load current divides equally between loops 1 and 2 and remains constant from t_4 to t_5. Current equal and opposite to the loop currents flows in the control circuit. Since the load current increases stepwise each half-cycle without bound, this circuit is unstable.

Case IV. Inductance and Resistance Load, Zero Control-Circuit Resistance

The zero control-circuit resistance may be realized, as discussed in case II. In this case, as shown in curves *D*, *E*, and *F* of Fig. 4, the current build-up in the interval t_1 to t_2 is of the form

$$i_L(t)=i_1(t_1)\epsilon^{-\frac{\omega}{Q}(t-t_1)}+\frac{E}{Z}\Big[\sin(\omega t-\tan^{-1}Q)$$
$$-\epsilon^{-\frac{\omega}{Q}(t-t_1)}\sin(\omega t_1-\tan^{-1}Q)\Big] \tag{7}$$

where $Q=\omega L/R$ and $Z^2=\omega^2L^2+R^2$. At t_2, the load current divides equally between the two loops but does not remain constant as in case III. Rather, it decays with a time constant $T=L/R=Q/\omega$ from t_2 to t_3 while the gates set and reset at normal voltage. At t_3, $i(t_3)=i(t_2)\epsilon^{-\frac{\omega}{Q}(t_3-t_2)}$ and the current buildup is again given by equation 7.

For $Q>0$, the load current at the voltage reversal point continues to increase each half-cycle until the transient increase during the conducting period is equal to the current decay during the previous half-cycle, that is

$$i_L[\omega t=(n+1)\pi]=i_L[\omega t=n\pi] \tag{8}$$

Let θ be the firing angle ($\theta/\omega=t_2-t_1$). Then the steady-state value of current at the voltage reversal point, obtained by combining equation 7 with conditions of equation 8, is

$$i_0=\frac{E}{Z(1-\epsilon^{-\pi/Q})}[\sin(\tan^{-1}Q)-\epsilon^{-\theta/Q}\sin$$
$$(\theta+\tan^{-1}Q)] \tag{9}$$

Since there is a definite value of output current and firing angle for each value of bias the circuit is stable.

Case V. Resistance and Inductance Load and Nonzero Control-Circuit Resistance

This case resembles the last two cases for the time interval t_0 to t_2; see Fig. 5. Gate 1 saturates at t_1 as gate 2 attains the bias-flux level. Loop 1 conducts and the load current is the same as in case IV. Loop 2 is blocked from further change by rectifier action. At t_2, the supply voltage reverses in sign and loop 2 has double set voltage. Now, however, the equivalent circuit is as illustrated in Fig. 6(A).

Gate 2 is assumed to be a perfect transformer and is replaced by the equivalent resistance $R_c=r_c/a^2$. The transformed set voltage of gate 2 causes a current i_2/a to flow in the control winding which, in turn, allows a current i_2 to flow in the gate 2 power winding. Then, for $i_2\leq i_1$

$$e=e_L=(i_1+i_2)R+L\frac{d}{dt}(i_1+i_2)\text{ and}$$
$$-e=i_2R_c+e_L \tag{10}$$

Hence

$$-2e=i_2R_c\text{ and }e=L\frac{di_L}{dt}+Ri_L \tag{11}$$

where $i_L=i_1+i_2$. For the condition $i_1=i_2$

$$i_L=2i_1=-4e/R_c \tag{12}$$

In the interval t_2 to t_3, gate 1 is still saturated and is considered a short circuit with the result that the load is still connected directly to the voltage supply, and load current continues to be expressed by equation 7. Loop 2 current increases as $i_2=2e/R_c$ to the point t_3, as shown in Figs. 5 and 6(B) ($i_1=i_2=1/2\ i_L$), when core 1 leaves the region of saturation. Beyond t_3, the load current remains evenly divided between the two loops because of the control-circuit coupling (see Appendix III) and decays as

$$i_L(t)=i_L(t_3)\epsilon^{-(t-t_3)/T};\ t_3<t<t_4 \tag{13}$$

where

$$T=\frac{L}{R+R_c/4}=\frac{L}{R+r_c/4a^2} \tag{14}$$

During the interval t_3 to t_4, gate 1 remains saturated. Gate 2 sets at a voltage, i.e., double voltage. At t_4, gate 1 leaves the saturated state and resets at normal voltage while gate 2 sets at normal voltage. At t_4, gate 2 saturates earlier in this half-cycle than did gate 1 during the preceding half-cycle because the interval of double set voltage from t_2 to t_3 has caused gate 2 to absorb the necessary amount of volt-seconds in a shorter interval. This change in the set interval caused by the double set voltage experienced during current commutation is a fundamental factor in producing instability. It should be noted that the double set voltage is caused by the load voltage adding set volt-seconds to the unsaturated core when the supply voltage reverses polarity and the load current starts to decline.

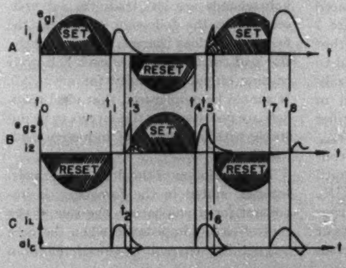
Gate 1 now acts as a perfect transformer and is replaced by the reflected control-circuit resistance R_c. Just before gate 2 saturates, the current in loop 1 is

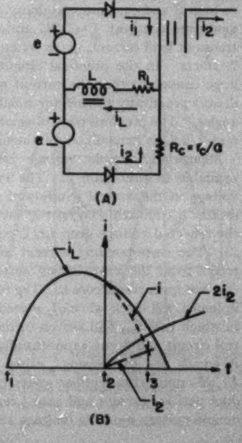
$$i_1(t_4)=\frac{2e(t_3)}{R_c}\epsilon^{-(t_4-t_3)/T};\ T=\frac{L}{R+R_c/4} \tag{15}$$

At t_4, double voltage is impressed across R_c and a current $i=2e(t_4)/R_c$ is possible in the outer loop, provided the rectifiers

Fig. 5 (below). Wave forms for case V, an inductance-resistance load and nonzero control-circuit resistance

Fig. 6 (right). Equivalent circuit and load current response during the interval in which load current flows, case V, an inductance-resistance load and nonzero control-circuit resistance





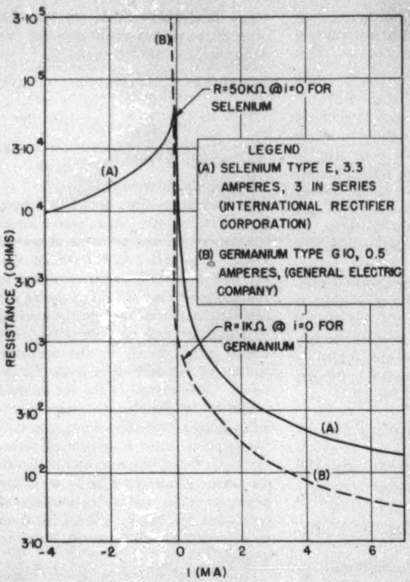


Fig. 7 (left). Rectifier d-c resistance

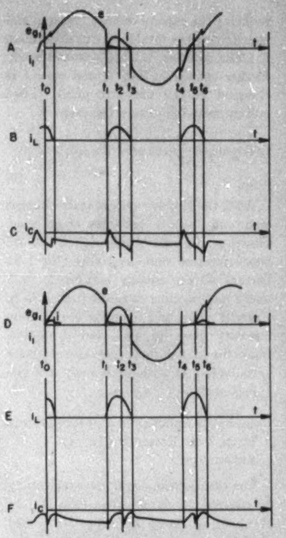


Fig. 8 (right). Wave forms from the experimental amplifier having inductive loading, finite control-circuit resistance, and either selenium (A, B, C) or germanium (D, E, F) rectifiers

do not block. To determine what loop 1 current does, consider the ratio

$$\frac{i}{i_1}(t_4^-)=\frac{e(t_4)}{e(t_3)\exp[-(t_4-t_3)/T]} \tag{16}$$

But $e(t_4)>e(t_3)$ from elementary consideration of the bias flux and gate voltage relations. Hence, $i>i_1$, and loop 1 current tends to become negative. Whatever load current remains at t_4 switches to loop 2 and continues to decay with a new time constant $T=L/R$, and a new transient load current, given by equation 7, starts. In the practical circuit, T is large enough so that the current quickly decays practically to zero for small firing angles. The previous control-circuit current $1/2\ i_L$ is immediately extinguished.

Meanwhile, double voltage becomes available to reset gate 1. The induced voltage in the control winding of gate 1 creates a circulating current i_c such that the reflected voltage drop $i_c r_c$ is equal to $2e$. The corresponding current ai_c subtracts from the magnetizing current of (the not quite reset) core 1. The rectifier in loop 1 will block when ai_c exceeds i_{mag}, in which case i_{mag} will switch to the control circuit and decay exponentially until it equals the (decreasing) value $2e/R_c$. At this point, the rectifier starts to conduct positive current and gate 1 resets at double voltage until the voltage reverses sign. If r_c is sufficiently large, the ai_c may not exceed i_{mag} and gate 1 resets at double voltage for the complete interval t_3 to t_4. Although effects caused by magnetizing currents may be discussed, they will be neglected in the analysis.

At t_5, the supply voltage changes sign again and loop 1 becomes conductive. The load current continues smoothly as before up to t_6 when the gate currents in the outside loop, increasing sinusoidally in the interval t_5 to t_6, become equal to $i_L/2$. During the interval t_5 to t_6, gate 1 sets at double voltage and gate 2 does not reset at all. The cycle continues as in the previous half-cycle, and the output current would eventually assume a constant value similar to case IV if the firing angle did not change. However, the reset volt-seconds are less than the set volt-seconds in the following cycle and the firing angle does change. Consequently, the load current increases each half-cycle causing circuit instability for any $Q>0$. It should be emphasized that ideal rectifiers are necessary for this circuit to be unstable with all degrees of inductive loading.

The restriction of the double reset volt-seconds, stated in the foregoing, is important for it emphasizes the role of the control-circuit impedance when the core magnetizing current is considered in a practical case. If, for example, R_c were very large, the discharge of the magnetizing current of gate 1 through R_c would cause a large transient voltage to appear across gate 1 with the result that the gate would continue to reset; the additional reset volt-seconds are a stabilizing influence. In the limit, if $R_c=\infty$, the circuit will be stable for any Q.

CASE VI. PRACTICAL CIRCUIT HAVING A RESISTIVE AND INDUCTIVE LOAD AND NONZERO CONTROL-CIRCUIT RESISTANCE

In any practical circuit, there are several complicating influences. Loop resistance decreases the voltage across the gate by its IR drop, and rectifier leakage allows some additional reset beyond the bias point. The actual magnetic and rectifier characteristics are much more complicated than the postulated ones. Rectifier capacitance has an appreciable effect.

Experimentally, the salf-saturating, full-wave magnetic amplifier with an RL load is stable in the region of low Q loads and high control-circuit resistance. It is interesting to note that, for most of the ideal cases previously discussed, instability results for any value of Q. It is the factors that are neglected in the idealized system that provide some degree of stability. These are the actual rectifier and core characteristics.

In the practical circuit, the magnetic characteristics consist of a continuous family of minor asymmetrical hysteresis loops, as indicated in Fig. 2(C). Rectifiers have resistance characteristics, as shown in Fig. 7. The operation of a circuit, considering some of these actual characteristics, is shown in Fig. 8. The voltage across gate 2 is no longer a sine wave when the rectifier characteristic is in a region of high resistance. This region occurs for currents less than a small, positive value; see Fig. 7. The rectifier current actually reverses during the last half of the reset interval t_0 to t_2. The rectifier resistance becomes large during this interval, and some of the reset voltage is consumed in rectifier voltage drop. The locus of ϕ versus 1 for gate 2 is one that cuts across the family of minor loops as the rectifier consumes its share of the voltage.

From t_1 to t_3, load current flows from loop 1. The forward rectifier resistance is added to the load resistance. At t_1, gate 1 conducts, causing double reset voltage to appear in loop 2. This double reset voltage divides between the back rectifier impedance and the reflected control-circuit resistance. The selenium rectifiers have appreciable capacitance so that, initially, the full double voltage appears across R_c, see t_1 in Fig. 8, curve *A*. This double voltage decreases rapidly as the rectifier capacitance quickly charges. At t_2, the supply voltage reverses sign and double set voltage is applied to loop 2. The rectifier resistance of loop 2 does not decrease immediately to a low value not only because of the slight positive current required to move the operating point on the rectifier characteristic to a region of low resistance, but also because the rectifier capacitance, effectively in parallel with the rectifier resistance, perpetuates the actual rectifier current in the negative direction beyond the voltage reversal point; see Appendix IV. When the rectifier current is sufficiently positive, the rectifier resistance decreases to its small forward value. Load current switching from loop 1 adds positive current to loop 2 and causes the rectifier resistance to decrease to a low value only slightly beyond the voltage reversal point. From t_2 to t_3, load current continues to flow in loop 1. At t_3, some load current switches to loop 2 (see case V) but the time constants with the selenium cells are so small that the oscillogram response is not discernible.

Germanium rectifiers have negligible capacitance, and the double loop voltage divides between the back rectifier resistance and the reflected control-circuit resistance. A galvanometer with a high-frequency response was used to record gate current of the amplifier employing germanium rectifiers. The load current that switched from one loop to the other is discernible in Fig. 8, curve *D*.

At t_3, the load current has decreased to zero, neglecting magnetizing current. Gate 1 leaves the saturated zone beyond t_3, and both gates absorb voltage at a normal rate. Magnetizing currents alone flow in both loops in the interval t_3 to t_4.

At t_4, gate 2 saturates at a different time in the half-cycle than the corresponding saturation time for gate 1 during the preceding half-cycle because of the intervals of modified set and reset voltages. During the interval t_4 to t_5, loop 1 has double reset voltage but the magnetizing current is in the region of high rectifier resistance, and voltage division occurs again. The circuit operation proceeds in this manner.

It is seen that, for other than perfect rectifiers, there are several opposing effects present. There is the interval of inhibited double set voltage during the commutation period caused by the set voltage division between forward rectifier resistance and reflected control-circuit resistance, and the interval of inhibited reset caused by the reset voltage division between the reflected control-circuit resistance and either back resistance of the germanium rectifier or back impedance of the selenium rectifier. In addition, there is the series *IR* drop which removes more set voltage than reset voltage because of the larger magnetizing currents during set: rectifier leakage, which does allow some reset beyond the bias point for sufficiently large reflected control resistance; and selenium rectifier capacitance which creates a lag in the application of double set voltage by holding the rectifier in a region of high resistance beyond the point of voltage reversal.

Analytical Determination of Stability Boundary

To analyze the practical circuit of case VI, it will be necessary to determine experimentally the initial conditions of operation and to show the exact basis of analysis. The particular full-wave, self-saturating magnetic amplifier used is described in Appendix I.

When the control-circuit resistance is adjusted to the critical value and the control current is slowly reduced, tiny load current pulses appear before the amplifier triggers; see Fig. 9. The firing angle for these pulses is about 20 degrees. In the analysis for the boundary of stability, the firing angle appears in the equations, and this value of 20 degrees is used in the calculations.

At the minimum output conditions, the gate voltage differs from the supply voltage. There is loss of reset volt-seconds caused by the *IR* drop in the forward loop resistance and the voltage wave forms are modified somewhat by the nonsinusoidal magnetizing currents. Whatever the exact situation may be, there does exist a balance between set and reset volt-seconds at the initial stable value of load current. The analysis will proceed from this point. The voltage wave forms will be restored to sinusoidal shapes to simplify the computations.

For small values of θ/Q, where θ is the firing angle, it is possible to estimate the commutation angle y; see Appendix II. This angle, which is the angle beyond the supply voltage reversal point for which load current persists, is given by

$$y=\theta\left(1-\frac{2\theta}{3Q}-\frac{4QR}{\theta R_c}\right) \qquad (17)$$

This equation is correct within 2 per cent for $\theta/Q<0.2$, and $R_c>500\ R$. For example, to ensure this accuracy for $\theta=20$ degrees, Q must be greater than 0.5.

With the circuit of case VI, shown in Fig. 10, the following are postulated: a constant forward rectifier resistance R_f, a constant back rectifier resistance R_b, and no rectifier capacitance. Let the first half-cycle be stable, so that

$$\theta_2^2=\theta_1^2\left(1-\frac{2\theta_1}{3Q}-\frac{4QR_1}{\theta_1 R_c}\right)^2$$
$$\cong\theta_1^2\left(1-\frac{4\theta_1}{3Q}-\frac{8QR_1}{\theta_1 R_c}\right) \qquad (18)$$

within the accuracy of the approximation, and $\theta_2=\theta_0$; see Fig. 11. R_1 is the total forward resistance of loop 1.

For the duration of θ_2, gate 1 sets at a voltage

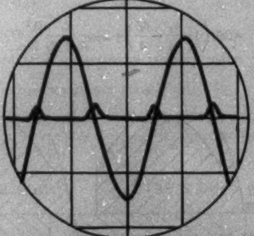


Fig. 9. Line voltage and minimum output current at the critically stable condition with a firing angle of about 20 degrees. Q=2.5. Zero current line not preserved. Current pulse=15 milliamperes peak

l, 74
1
55

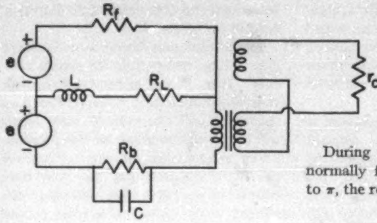


Fig. 10. Equivalent circuit for one phase of operation in case VI

$$e_1=(2e-i_2R_f)\frac{R_c}{R_c+R_f}\backsimeq 2e\frac{R_c}{R_c+R_f} \qquad (19)$$

The current i_2 decreases during the commutation period because some current switches to loop 1, and the load current decays. Consequently, i_2R_f remains small throughout the interval and it will be neglected in the analysis.

Gate 2 does not reset at all during this interval. The change of flux in the gate for an applied voltage of $e=E\sin\omega t$ is

$$\Delta\phi=\frac{1}{N}\int_{t_1}^{t_2}E\sin\omega t\,dt=\frac{1}{\omega N}\int_{x_1}^{x_2}E\sin dx \qquad (20)$$

$$=\frac{E}{\omega N}(\cos x_1-\cos x_2) \qquad (21)$$

Let $E=N\omega\phi_s$, the usual operating condition. Then

$$\Delta\phi=\phi_s(\cos x_1-\cos x_2) \qquad (22)$$

For $x_1=\theta_2$, then $x_2=\pi-\theta_3$. When θ_2 and θ_3 are small, the change in flux from θ_2 to $\pi-\theta_3$ is

$$\Delta\phi=2\phi_s\left(1-\frac{\theta_2^2}{4}-\frac{\theta_3^2}{4}\right) \qquad (23)$$

The complete change of flux in gate 1 during the interval 0 to π is

$$\Delta\phi_1=2\phi_s\left(1-\frac{\theta_2^2}{4}-\frac{\theta_3^2}{4}\right)+$$
$$2\phi_s\frac{R_c}{R_c+R_f}(\cos\theta-\cos\theta_2) \qquad (24)$$
$$=2\phi_s\left(1-\frac{\theta_2^2}{4}-\frac{\theta_3^2}{4}\right)+2\phi_s\frac{R_c}{R_c+R_f}\frac{\theta_2^2}{2}$$

During this half-cycle, gate 2 resets normally from θ_2 to $\pi-\theta_3$. From $\pi-\theta_3$ to π, the reset voltage is

$$-e_2=(2e-i_1R_f)\frac{R_c}{R_c+R_b}\backsimeq 2e\frac{R_c}{R_c+R_b} \qquad (25)$$

The change in flux of gate 2 is

$$-\Delta\phi_2=\phi_s[\cos\theta_2-\cos(\pi-\theta_3)]+2\phi_s[\cos(\pi-\theta_3)-\cos\pi]\frac{R_c}{R_c+R_b}$$
$$=2\phi_s\left(1-\frac{\theta_2^2}{4}-\frac{\theta_3^2}{4}\right)+2\phi_s\frac{R_c}{R_c+R_b}\frac{\theta_3^2}{2} \qquad (26)$$

The boundaries of stability are determined by

$$\Delta\phi_1=-\Delta\phi_2 \text{ and } \theta_1=\theta_3 \qquad (27)$$

Substituting equations 18, 24, and 26 into the conditions of equation 27, the following conditional equation for stability is obtained

$$2\phi_s\left(1-\frac{\theta_2^2}{4}-\frac{\theta_3^2}{4}\right)+2\phi_s\frac{R_c}{R_c+R_f}\frac{\theta_1^2}{2}\times$$
$$\left(1-\frac{4\theta_1}{3Q}-\frac{8QR_1}{\theta_1R_c}\right)=2\phi_s\left(1-\frac{\theta_2^2}{4}-\frac{\theta_3^2}{4}\right)+$$
$$2\phi_s\frac{R_c}{R_c+R_b}\frac{\theta_3^2}{2} \qquad (28)$$

The terms involving the interval of normal set and reset, θ_2 to $\pi-\theta_3$, cancel and the flux balance reduces to that of equating set volt-seconds of gate 1 during the commutation period to the reset volt-seconds of gate 2 during the normal conduction period. Finally

$$\frac{R_c}{R_c+R_b}=\frac{R_c}{R_c+R_f}\left(1-\frac{4\theta_1}{3Q}-\frac{8QR_1}{\theta_1R_c}\right) \qquad (29)$$

Solving

$$R_c=\frac{\left(1-\frac{4\theta_1}{3Q}-\frac{8QR_1}{\theta_1R_c}\right)R_b-R_f}{\frac{4\theta_1}{3Q}+\frac{8QR_1}{\theta_1R_c}} \qquad (30)$$

In general, R_f is negligible with respect to the R_b term, and equation 30 is approximately

$$R_c=\left(\frac{3Q}{4\theta_1}-1\right)R_b-18\frac{Q^2R_1}{\theta_1^2(3Q-4\theta_1)};\ 3Q>4\theta_1 \qquad (31)$$

The R_b term is quite large with respect to the R_1 term for $3Q$, more than 1 per cent greater than $4\theta_1$, and equation 31 simplifies to

$$R_c=\left(\frac{3Q}{4\theta_1}-1\right)R_b;\ 3Q>4.04\theta_1 \qquad (32)$$

To derive equation 29, the assumption has been made for

$$2e-i_2R_f=2e(1-\gamma_1) \text{ and } 2e-i_LR_f=2e(1-\gamma_2) \qquad (33)$$

that

$$1-\gamma_1\backsimeq 1-\gamma_2 \qquad (34)$$

Since the forward rectifier voltage drop is relatively small compared with $2e$ and the difference in equation 34 is smaller, the assumption will not introduce a large error.

Equation 29 has been derived for idealized core materials with zero magnetizing current, i.e., a vertical curve of magnetic induction versus magnetic intensity. This assumption is valid for small to moderate values of R_c, but breaks down completely as R_c approaches infinity. Then, the decay of the magnetizing current allows gate 2 to reset as a voltage not greater than $2e$. The error introduced by neglecting the reset effects of the magnetizing current will result in a conservative prediction of the stability border.

The load current which is divided between the two loops at θ_3 decays with a time constant $T=L/(R_L+R_c/4)$. It is shown in Appendix II that the flux change caused by the load currents adds an equal amount of set and reset flux to

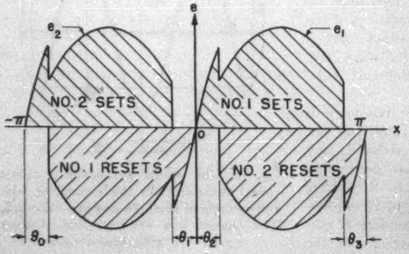


Fig. 11 (left). Voltage wave forms for analysis of case VI showing maximum possible set and reset areas

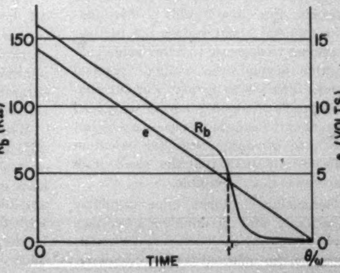


Fig. 12 (right). Germanium rectifier resistance as a function of the voltage $e=42.5\sin(\theta-\omega t)$

ol. 74
. 1
955

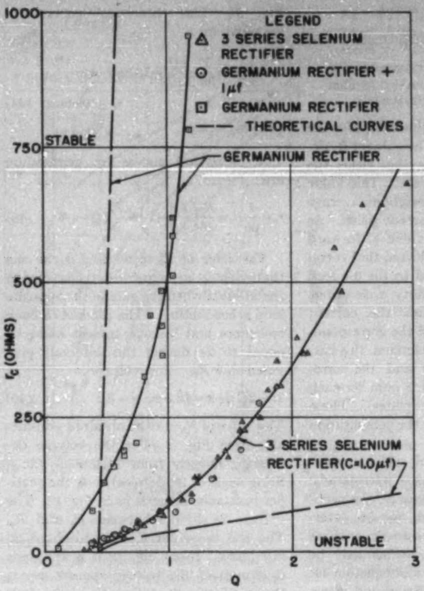


Fig. 13. Boundary of stability

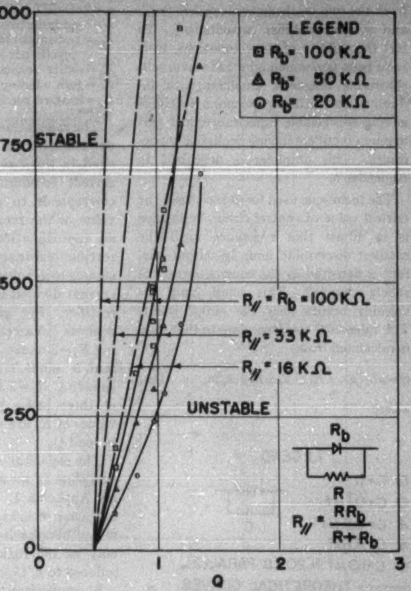


Fig. 14. Effect on stability boundary of shunting germanium rectifiers with resistors

gates 1 and 2 respectively. Consequently, this flux change does not influence the stability equation 28 because it merely adds identical terms to both sides of the equation.

Effect of Rectifier Capacitance

Next, consider the effect of rectifier capacitance. The equivalent circuit for the interval $\pi-\theta_3$ to π is shown in Fig. 12. Gate 1 is saturated and gate 2 is resetting. Gate 2 voltage in this interval (see Appendix IV) is

$$e_2=2E[\alpha(\omega t-\theta_3)-\theta_3(1-\alpha)\epsilon^{-t/\alpha\tau}+\alpha\omega\tau(1-\alpha)(1-\epsilon^{-t/\alpha\tau})] \quad (35)$$

where ωt is measured from $\pi-\theta_2$ and

$$\alpha=\frac{R_c}{R_c+R_b} \text{ and } \tau=R_b c$$

The first term is the steady-state proportional term for small firing angles; the second, the initial transient; the third, the steady-state lag with its transient term. The initial transient provides additional reset while the steady-state lag term adds set voltage. At the voltage reversal point

$$e_3(\theta/\omega)=2E(1-\alpha)[\alpha\omega\tau-(\alpha\omega\tau+\theta)\epsilon^{-\theta/\alpha\omega\tau}] \quad (36)$$

For $Q>0.5$, the experimental value of the exponent is small, and

$$e_3(\theta/\omega)=2E(1-\alpha)\frac{\theta^2}{2\alpha\omega\tau} \quad (37)$$

The gate voltage at $x=\pi$ is positive under these conditions, a result in agreement with Fig. 8.

The analysis of Appendix IV indicates that the current through the rectifier resistance (in the equivalent circuit) is perpetuated in the reverse direction beyond the voltage reversal point. At this reversal point, however, load current starts to switch to the loop under consideration, quickly reversing the rectifier current, and returns the rectifier resistance to the low, forward value. The influence of rectifier capacitance on reset flux is largely confined to the interval $\pi-\theta_3$ to π. The reset flux during this interval obtained by integration of the complete voltage equation between these limits is

$$\Delta\phi_2=\frac{2E}{N}\left\{\frac{\alpha\theta^2}{2\omega}-\theta^2\frac{\alpha}{\omega}+\alpha\omega\tau(1-\alpha)\left[\frac{\theta}{\omega}+\alpha\omega\tau(\epsilon^{-\pi/\alpha\omega\tau}-1)+\theta(1-\alpha)\times(\epsilon^{-\theta/\alpha\omega\tau}-1)\alpha\tau\right]\right\} \quad (38)$$

where θ has been written for θ_3.

For the experimental amplifier, the exponential term may be approximated to 1 per cent by its power series expansion for $R_c>360$ ohms, and

$$\Delta\theta_3=\frac{2F}{N\omega}\frac{\theta_3^3}{2}\left(1-\frac{2}{3}\frac{1-\alpha}{\alpha\omega\tau}\theta_3\right) \quad (39)$$

Performing the flux balance as before

$$R_c=\frac{R_b-R_f/\theta}{\alpha\omega\tau}Q \quad (40)$$

The R_f term is generally negligible with respect to the R_b term, and the equation becomes

$$R_c=\frac{R_b}{2\omega\tau}Q=\frac{1}{2\omega C}Q=\frac{n}{2\omega C_1}Q \quad (41)$$

where

n = number of selenium rectifier plates in series
C_1 = capacitance per plate

For the special case of $R_c=\alpha=0$, equation 38 reduces to $\Delta\phi_2=0$, and, for this case

$$R_c=0;\ Q=\frac{4}{3}\theta_1 \quad (42)$$

Experimental Determination of the Stability Boundary

In the preceding sections, equations for the stability region have been derived from the two mathematical models (with and without rectifier capacitance). To test the validity of these equations, tests have been performed on a full-wave, self-saturating magnetic amplifier with inductive load utilizing selenium rectifiers having appreciable capacitance, and germanium rectifiers having negligible capacitance. This amplifier is described in Appendix I.

The technique used for determining the critical value of control-circuit resistance is to adjust that resistance until the smallest discernible jump in output current is detected as the control current is slowly changed. This small jump invariably occurs near the cutoff point. The value of Q corresponding to that load is calculated from

$$Q=\omega L/(R_{ac}+R_{L}+R_f+R_s+R_g) \qquad (43)$$

where

$\omega=2\pi f=377$ radians per second
L = load inductance
R_{ac} = resistance of load inductance at 60 cycles per second
R_s = source resistance = 0.05 ohm
R_g = gate winding resistance = 0.78 ohm
R_f = forward rectifier resistance

The calculation is routine except for the determination of R_f. The value of R_f desired in this case is one that yields the correct commutation time. This value corresponds to some weighted average value of the rectifier current pulse. As an approximation, the value of forward rectifier resistance is selected that corresponds to a current equal to the d-c load current divided by the duty cycle of one rectifier. For germanium, this current is about 0.5 ampere, and the corresponding $R_f=2$ ohms. For selenium, this current is about 1 ampere and the corresponding $R_f=1$ ohm (1/3 ohm for each of three series rectifier plates). These values of R_f are used in the computation of load Q.

The capacitance of each selenium rectifier plate is measured as 3 microfarads; see Appendix I. Fortunately, the rectifier time constant need not be determined because it enters into the analysis only as the ratio R_b/τ, which may be reduced to n/C_1, as shown in equation 41.

A comparison of predicted and measured system performance is shown in Fig. 13. The theoretical curves for selenium are

$$r_c=\frac{1}{36}R_c=\frac{1}{36}\,\frac{nQ}{2\omega C_1}=12.3nQ\text{ ohms};\qquad r_c\geq 10\text{ ohms} \qquad (44)$$

$$Q=4/9;\quad r_c=0$$

The theoretical curves for germanium are

$$r_c=\frac{1}{36}R_c=\frac{R_b}{36}\left(\frac{3}{4}\,\frac{\theta}{\theta}-1\right)=\frac{R_b}{16}(Q-4/9) \qquad (45)$$

The value of R_b to be used is the one that yields the actual reset flux during the interval when the gate in the opposite loop is conducting. The product of back resistance and leakage current must be equal to $2e$ during this interval, provided $R_b \gg R_c$. This voltage is

$$2e=2E\sin x=42.5\sin x=R_b i \qquad (46)$$

The value of R_b may be obtained graphically from Fig. 7. For the voltage decreasing linearly from 14.4 volts (at a firing angle of 20 degrees) to 0, the rectifier resistance behaves as in Fig. 12. The voltage $2e$ divides between R_b and R_c. The test range for R_c runs to about 20 kilo-ohms. From Fig. 12 it is apparent that most of the voltage appears across the rectifier in the interval 0 to t'. It is reasonable, then, to use an average value of R_b in this interval (10,000 ohms) as a first approximation. During the interval t' to θ/ω, nearly all the voltage ap-

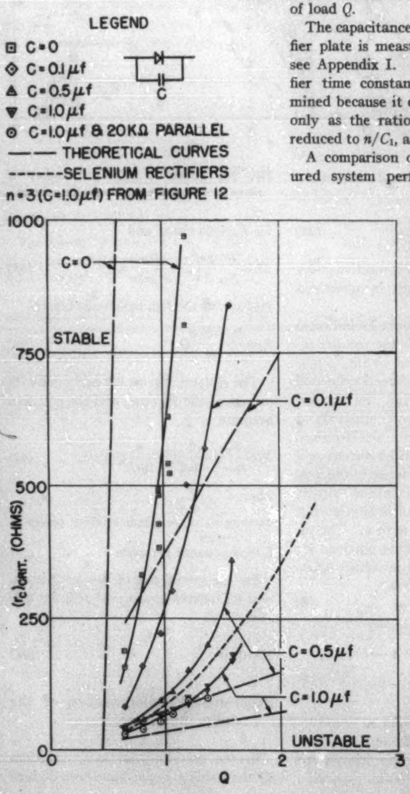


Fig. 15 (left). Effect on stability boundary of shunting germanium rectifiers with capacitors

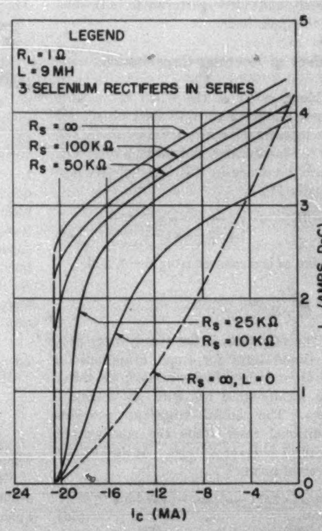


Fig. 16 (right). Effect on transfer curves of shunting selenium rectifiers with resistors R_s

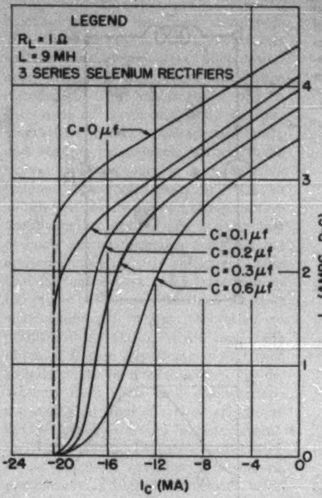


Fig. 17 (left). Effect on transfer curves of shunting selenium rectifiers with capacitors C.

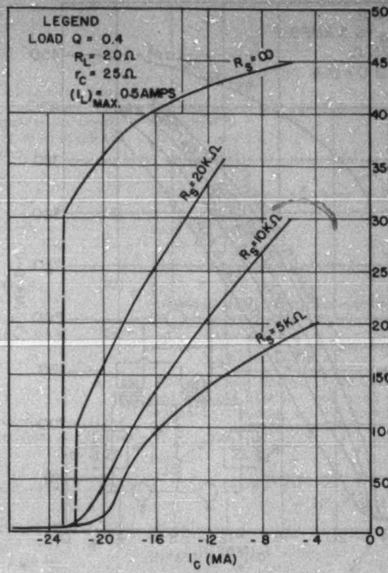


Fig. 18 (right). Effect on transfer curves of shunting germanium rectifiers with resistors R_s

pears across the core and adds reset volt-seconds. The voltage in this interval is quite low, and the additional reset flux which is a conservative influence in the predicted stability boundary will be neglected.

With this value of R_b, the theoretical equation for the case with germanium rectifiers becomes

$$r_c = 6{,}250(Q - 0.44) \tag{47}$$

Equation 47 indicates stability for $Q > 0.44$. Fig. 13 shows the tendency of the amplifier to behave in this manner. The predicted boundary for this case is on the conservative side.

With selenium rectifiers, the rectifier capacitance dominates the system and provides a greater region of stable operation. In the range investigated, there is always a stable operating point for any value of Q for a sufficiently large r_c. The predicted stability region for this case is not on the conservative side.

Effect of Rectifier Shunting

To verify the form of the analytical equations, a series of stability tests on the amplifier using shunted germanium rectifiers was performed. Fig. 14 shows the effect of shunting the rectifiers with resistors, which modifies the value of R_b in the equation. It is seen that the variation in the stability boundary behaves in the predicted manner and that the agreement between theory and experiment improves as the shunting resistor becomes smaller. Resistance shunts not only improve stability but also minimize the nonlinearity and environmental effects of the rectifiers.

Fig. 15 shows the effect of shunting the germanium rectifiers with capacitors. The analytical expression involving C now applies, and the boundary of stability does behave in the predicted manner. The theory also predicts that back rectifier resistance has negligible effect on the stability boundary within the limits of the simplifying assumptions when rectifier capacitance is appreciable. The dotted curve in Fig. 15 is for selenium rectifiers having 1-microfarad capacitance. It does correspond closely with the 1-microfarad shunting-capacitor curve for germanium even though the back resistance of germanium is many times that of selenium.

An unstable amplifier may be stabilized by shunting the rectifiers by resistance, capacitance, or both. A series of transfer curves is presented in Fig. 16, showing the stabilizing effect of shunting selenium rectifiers with resistors. Fig. 17 shows the same effect using capacitors to shunt selenium rectifiers. Figs. 18 and 19 show the same effects on the transfer curves for an amplifier employing germanium rectifiers.

Discussion of Results

No attempt was made in the analysis to include the effects of magnetizing current on stability, although a qualitative discussion of some of these effects is presented. The result of the omission is a conservative prediction of the stability boundary. A further simplification introduced in the analysis is the assumption of a constant value for back rectifier resistance. For a rectifier without appreciable capacitance, this back resistance is a primary factor in the determination of the stability boundary. The fact that this back resistance is far from constant in the regions in which its value is a key factor is a strong recommendation for its more accurate representation in future analyses.

The value of $\theta_f = 20$ degrees is used throughout the calculation for the stability borders, a value based on experimental observation and measurement. Any convenient firing angle could have been used as the maximum stable angle in the analysis but it was noted that output current was completely insignificant for angles less than 20 degrees. Although an amplifier may have an unstable jump[1] for firing angles less than 20 degrees, the results are of little practical importance unless the firing angle exceeds 20 degrees.

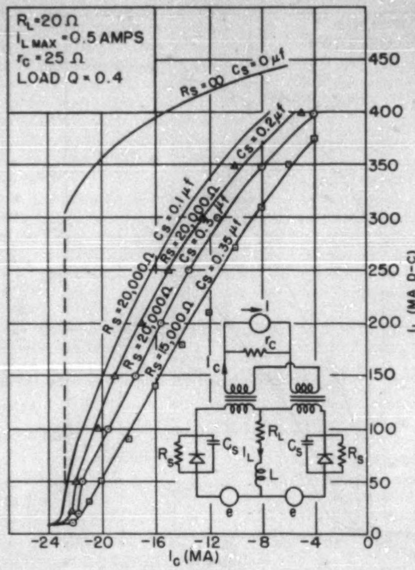


Fig. 19 (left). Effect on transfer curves of shunting germanium rectifiers with parallel resistance and capacitance

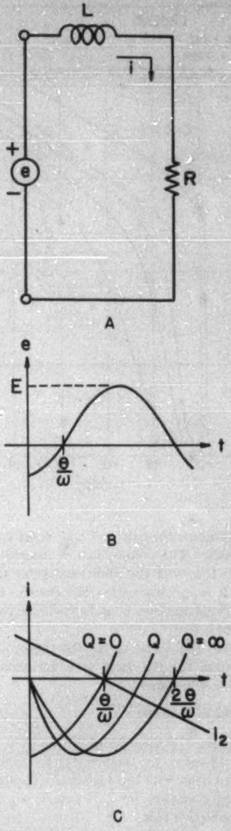


Fig. 20 (right). Equivalent circuit and wave forms for the determination of the commutation period

This value of 20 degrees, then, represents a lower limit on the small-angle analysis.

Within the framework of these approximations and the idealization of the magnetic characteristics, the boundary of stability for the full-wave, self-saturating magnetic amplifier with inductive load has been derived. For the amplifier with selenium rectifiers, (appreciable rectifier capacitance) the stability boundary is given by

$$R_c = \frac{Q}{2\omega C} \qquad (48)$$

For germanium rectifiers (negligible rectifier capacitance)

$$R_c = R_b\left(\frac{3Q}{4\theta_f} - 1\right) \qquad (49)$$

The analysis is not accurate enough to justify additional terms in equations 48 and 49.

It may be deduced from equation 49 that, for an amplifier having R_c slightly less than the critical value, θ_f and the corresponding output current will increase only to the stable value of θ_f calculable from equation 49. The circuit is not self-regenerating in that full output, i.e., $\theta_f = 180$ degrees, is not achieved for any R_c less than the critical one. This observation is confirmed experimentally. The amount of unstable jump in output may be controlled by adjusting R_c. The same behavior is exhibited by the amplifier with selenium rectifiers, although equation 48 does not predict this effect.

A further check upon the validity of the form of equation 49 has been obtained by increasing the a-c supply voltage $e = E \sin \omega t$. Equation 49 predicts that large values of the initial firing angle θ_f improve the stability border by allowing smaller values of R_c. If the supply voltage is increased, the gates are no longer able to absorb the full half-cycle of voltage and, consequently, they must fire before the end of each half-cycle, regardless of the amount of bias. As the voltage is increased, θ_f must increase. An experimental check of the critical R_c with constant R_L and Q has confirmed the predicted decrease in critical R_c with increasing voltage.

This paper describes several methods of obtaining high stable gain with inductive loading when negative feedback is introduced. Examples of the influence of negative feedback obtained by rectifier shunting are illustrated in Figs. 14 through 19. An incidental advantage of negative feedback by rectifier shunting is a relative independence of amplifier performance on varying rectifier characteristics.

Summary

This paper may be considered as a first deviation from the case of idealized components to predict the performance of magnetic amplifiers. Although the idealization of components is convenient for determining the gross aspects of magnetic amplifier performance, it has been shown here and elsewhere that the departures from the ideal determine performance in many instances.[7,8] Intimate knowledge of the effects of these departures will enable the designer to minimize them in some cases and utilize them to advantage in others.

For the full-wave, self-saturating magnetic amplifier, a relation between the constraint of the control circuit and the reactive loading has been developed. The effects of the rectifier imperfections such as forward resistance, back resistance, and back capacitance have been introduced. The imperfections of the gate such as hystereses, eddy current, magnetizing current, leakage reactance, and winding capacitance, at best, have been introduced only in a qualitative fashion. In a recent paper[9] it has been shown that the rate of flux change within the magnetic amplifier gate is a function of many factors including core material and core material strip thickness. It is thought that, even at the low frequencies (60 cycles per second) at which the tests were conducted, better theoretical-experimental correlation would have been obtained if gate imperfections had been included. At higher frequencies, the deviations from ideal gate characteristic are such that their effects on the amplifier performance would become increasingly important. Further detailed studies of magnetic amplifier performance, particularly at higher frequencies, could profitably explore these characteristics of the gate along with a more detailed analysis of the dynamic back resistance of the rectifier.

Appendix I. Description of Experimental Amplifier

The Deltamax test cores are tape-wound toroids, Arnold Engineering Company, case no. *4180-D1*. Core size: outside diameter = 3.00 inches, inside diameter = 2.50 inches, height $h=0.50$ inch. Tape size $t=0.001$ inch. Gate winding = 600 turns and 0.78 ohm each, and control winding = 100 turns and 0.66 ohm each.

The selenium rectifiers are International Rectifier Corporation type-*E* rated at 3.3 amperes d-c. They have 3-microfarad capacity per plate as measured on a *GR* bridge at 1,000 cycles per second. The germanium rectifiers are General Electric Company type-*G10*. The d-c resistance characteristics of both rectifiers are shown in Fig. 7.

Appendix II. Computation of Load Current

The interval beyond the voltage reversal point during which load current continues to flow is known as the commutation period. Such an interval exists for an inductive load in a full-wave, self-saturating magnetic amplifier. For the idealized system containing perfect rectifiers and negligible magnetizing current, the commutation period may be evaluated.

Consider the series *RL* circuit having an impressed voltage; see Fig. 20(A) and (B)

$$e=0;\ t<0$$

$$e=E\sin(\omega t-\theta);\ t>0 \qquad (50)$$

The loop equation is

$$e=L\frac{di}{dt}+Ri;\ i(0)=0 \qquad (51)$$

First consider the case in which no load current switches from one loop to the other during the complete commutation period. To find the time t_1 (L, R, θ) at which the current first returns to zero, let

$$\omega t-\theta=x;\ Q=1/p=\omega L/R;$$

$$\frac{d(\ )}{dt}=\frac{d(\ )}{dx}\frac{dx}{dt}=\omega\frac{d(\ )}{dx} \qquad (52)$$

then

$$E\sin x=\omega L\frac{di}{dx}+Ri;\ i(-\theta)=0 \qquad (53)$$

$$\frac{E}{R}\sin x=Q\frac{di}{dt}+i=\frac{E}{Z}\sqrt{1+Q^2}\sin x \qquad (54)$$

where

$$Z^2=R^2+\omega^2L^2=R^2(1+Q^2)$$

Solutions of $i(x)$ for three values of Q are shown in Fig. 20(C) for constant Z. Equation 54 may be written

$$\frac{di}{dx}+pi=\frac{E}{Z}\sqrt{1+p^2}\sin x;\ i(-\theta)=0 \qquad (55)$$

The solution to equation 55 is

$$i(x)=\frac{E}{\sqrt{1+p^2}}[\epsilon^{-p(x+\theta)}(p\sin\theta+\cos\theta)+p\sin x-\cos x] \qquad (56)$$

as can be verified by substituting into the differential equation and the initial condition equation.

For $x=y$, the current is again zero, and the equation for y is

$$\epsilon^{-p(y+\theta)}(p\sin\theta+\cos\theta)=\cos y-p\sin y \qquad (57)$$

Rearranging

$$\epsilon^{-p\theta}(p\sin\theta+\cos\theta)=\epsilon^{+py}(\cos y-p\sin y) \qquad (58)$$

The desired solution to this equation is the smallest value of time $t_1=y/\omega$ for which it is satisfied.

For the special case of small commutation periods and large circuit Q, equation 58 may be written as a finite power series expansion of the component parts as follows

$$\left(1-p\theta+\frac{1}{2}p^2\theta^2-\frac{1}{6}p^3\theta^3+\frac{1}{24}p^4\theta^4-\ldots\right)\times\left(p\theta-\frac{1}{6}p\theta^3+\frac{1}{120}p\theta^5-\ldots+1-\frac{1}{2}\theta^2+\frac{1}{24}\theta^4-\ldots\right)=\left(1+py+\frac{1}{2}p^2y^2+\frac{1}{6}p^3y^3+\frac{1}{24}p^4\theta^4+\ldots\right)\left(1-\frac{1}{2}y^2+\frac{1}{24}y^4-\ldots-py+\frac{1}{6}py^3-\frac{1}{120}py^5+\ldots\right) \qquad (59)$$

Simplifying equation 59, and keeping terms through y^4 yields

$$y^2=\theta^2\frac{1-\frac{2}{3}p\theta+\frac{1}{12}\theta^2(3p^2-1)}{1+\frac{2}{3}py+\frac{1}{12}y^2(3p^2-1)} \qquad (60)$$

The solution of this equation accurate to terms in θ^2 is

$$y=\theta\left(1-\frac{2}{3}p\theta\right) \qquad (61)$$

Iterating equation 60 yields

$$y=\theta\left(1-\frac{2}{3}p\theta+\frac{4}{9}p^2\theta^2\right) \qquad (62)$$

good for terms including θ^3. For

$$\frac{4}{9}p^2\theta^2<0.02\ (\text{i.e., }\theta^\circ<12Q);\ \omega t_1=y=\theta\times\left(1-\frac{2}{3}p\theta\right)\pm2\text{ per cent} \qquad (63)$$

Next consider the case in which load current switches from the loop under consideration to the other in the amount $i_2=2e/R_c$. To find the time t_2 (L, R, R_c) at which $i_2=i/2$, the equation of condition 57 becomes

$$\epsilon^{-p\theta}(p\sin\theta+\cos\theta)=\epsilon^{py}(\cos y-(p-d)\sin y) \qquad (64)$$

where

$$d=4Z\sqrt{1+p^2}/R_c$$

For the special case of small commutation periods, equation 64 may be expanded in a finite power series as before, and the commutation angle becomes

$$y=\theta\left(1-\frac{2}{3}p\theta-\frac{4R}{p\theta R_c}\right)\pm2\text{ per cent};\quad p\theta<0.2,\ R_c>500R \qquad (65)$$

Appendix III. Circuit Analysis: Both Loops Conducting and Neither Gate Saturated

This circuit configuration may occur in the case of the full-wave, self-saturating magnetic amplifier with inductive load, provided that the two loop currents are equal (and are half the load current) and the control current is equal to the negative of either loop current times a; see Fig. 1. Then

$$e=e_L+e_1 \qquad -r_ci_c=a(e_1+e_2)$$

$$-e=e_L+e_2 \qquad e_L=R_Li_L+L\frac{di_L}{dt} \qquad (66)$$

Let

$$i=i_1=i_2=\frac{1}{2}i_L=-ai_c \qquad (67)$$

Then

$$e_1+e_2=-2e_L=r_ci/a^2\text{ and }e_L=-i_Lr_c/4a^2 \qquad (68)$$

Also

$$L\frac{di_L}{dt}+\left(R_L+\frac{r_c}{4a^2}\right)i_L=0 \qquad (69)$$

The solution of equation 69 is

$$i_L(t)=i_L(0)\epsilon^{-t/T}$$

where

$$T=L/(R_L+r_c/4a^2) \tag{70}$$

and $i_L(0)$ is the current at $t=0$. In addition, gate 1 sets and gate 2 resets at an increased voltage

$$e_1=-e_2=e+i_Lr_c/4a^2=e+i_L(0)\frac{r_c}{4a^2}\epsilon^{-t/T} \tag{71}$$

The additional set or reset flux from the increased voltage over an interval 0 to π/ω is

$$\Delta\phi=\frac{1}{N}\int_0^{\pi/\omega}\frac{r^2}{4a^2}\epsilon^{-t/T}dt\approx\frac{Tr_c}{4a^2N}i_L(0) \text{ for } T\ll\pi/\omega \tag{72}$$

$$\Delta\phi\approx2\phi_s\frac{\omega Tr_c}{8Ea^2}i_L(0)$$

In any practical amplifier $T\approx4a^2L/r_c$, and equation 72 becomes

$$\Delta\phi\approx2\phi_s\frac{\omega iL_L(0)}{2E} \tag{73}$$

The time $t=0$ in this analysis corresponds to the time t_2 in Fig. 6. For the condition of impending instability, the firing angle $\theta_f=20$ degrees. The current cannot exceed $(2Ea^2/r_c)\sin\theta_f$. Then

$$\Delta\phi\approx2\phi_s\frac{\omega L}{27r_c} \tag{74}$$

For the range of experimental data, $\Delta\phi/2\phi_s$ ranges from 0.1 to 2 per cent.

The condition of impending instability is defined by equal set and reset flux over the half-cycle. The increment of flux change adds equal amounts of set and reset flux to the two gates. The magnitude of the additional flux change is small so that the principle of superposition may be applied. Consequently, the stability effects on both cores cancel and there is no resulting effect on stability.

Appendix IV. Effect of Rectifier Capacitance on Reset Flux

In a full-wave, self-saturating magnetic amplifier with inductive load, the load current flows in a series of pulses, each pulse starting at the firing angle and continuing through the voltage reversal point to the end of the commutation period. During this interval, rectifier capacitance has its greatest influence on reset flux. Not only does it modify the reset voltage of the gate but it also tends to keep the rectifier in a region of high resistance by perpetuating the rectifier current in the negative direction beyond the voltage reversal point. At this reversal point, however, some of the load current switches to the loop being considered and causes the rectifier resistance to drop to a low value only slightly beyond the reversal point.

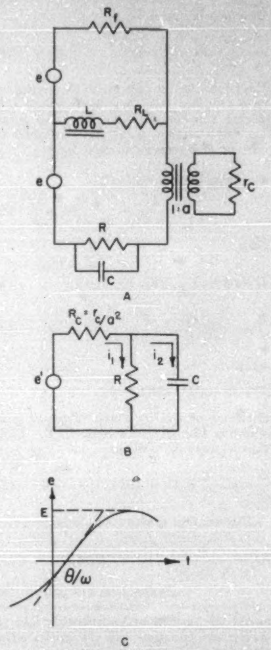


Fig. 21. Equivalent circuits and voltage wave forms for the determination of the effect on reset flux of rectifier capacitance during the flow of load current

To evaluate these effects on the reset flux, the equivalent circuit of Fig. 21(A) and its approximation (B) must be considered. The voltage impressed in this circuit, which (ideally) is double reset voltage, see Fig. 21(C), is

$$e'=e+\frac{Z_L}{Z_L+R_f}e=2e\left(1-\frac{1}{2}\,\frac{R_f}{R_f+Z_L}\right)=2e(1-\gamma) \tag{75}$$

In the analysis, γ is neglected; see equation 34. The voltage e is

$$e=0;\ t<0$$

$$e=E\sin(\omega t-\theta);\ t>0 \tag{76}$$

For $\theta\ll\pi$, the voltage e may be approximated, dotted line of Fig. 21(C), by

$$e=E(\omega t-\theta);\ 0<t<2\omega/\theta \tag{77}$$

The gate voltage is

$$e_{R_c}=2e[\alpha(\omega t-\theta)-\theta(1-\alpha)\epsilon^{-t/\alpha\tau}+\alpha\omega\tau(1-\alpha)(1-\epsilon^{-t/\alpha\tau})] \tag{78}$$

In the interval 0 to θ/ω, the flux change becomes

$$\Delta\phi=\frac{1}{N}\int_0^{\theta/\omega}e_{R_c}dt=\frac{2E}{N}\left[\left(\frac{\alpha\theta^2}{2\omega}-\frac{\alpha\theta^2}{\omega}\right)+\alpha\omega\tau(1-\alpha)\left(\frac{\theta}{\omega}+\alpha\tau\epsilon^{-\theta/\alpha\omega\tau}-1\right)+\theta(1-\alpha)\alpha\tau(\epsilon^{-\theta/\alpha\omega\tau}-1)\right] \tag{79}$$

For $\theta/\alpha\omega\tau$ small, the exponentials may be approximated by finite power series and through terms in θ^3. Then, equation 79 becomes

$$\Delta\phi=-\frac{2E}{N\omega}\,\frac{\theta^2}{2}\left(1-\frac{2}{3}\,\frac{1-\alpha}{\alpha\omega\tau}\theta\right) \tag{80}$$

References

1. An Instability of Self-Saturating Magnetic Amplifiers Using Rectangular Loop Core Materials, **S. B. Batdorf, W. N. Johnson.** *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 223–28.

2. Steady-State Analysis of Self-Saturating Magnetic Amplifiers Based on Linear Approximations of the Magnetization Curve, **W. H. Esselman.** *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 451–59.

3. Some General Properties of Magnetic Amplifiers, **J. M. Manley.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 39, 1951, p. 243.

4. Series-Connected Magnetic Amplifier with Inductive Loading, **T. G. Wilson.** *AIEE Transactions*, vol. 71, pt. I, Jan. 1952, pp. 101–10.

5. Saturable Reactors with Inductive D-C Load, **H. F. Storm.** *AIEE Transactions*, vol. 71, pt. I, Nov. 1952, pp. 335–43.

6. Magnetic Amplifiers with Inductive D-C Load, **H. F. Storm.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 467–75.

7. The Effective Feedback Ratio of Magnetic Amplifiers, **L. A. Finzi, G. F. Pittman, Jr., H. L. Durand.** *AIEE Transactions*, vol. 71, pt. I, Apr. 1952, pp. 157–64.

8. Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and Core Functions, **M. Frank, S. Rabotnick, J. R. Walker.** *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 182–88.

9. Flux Resetting Characteristics of Several Magnetic Materials, **H. Huhta.** *Ibid.*, pp. 111–14.

The Ring Modulator as a Polarized Rectifier

A. J. HERMONT
MEMBER AIEE

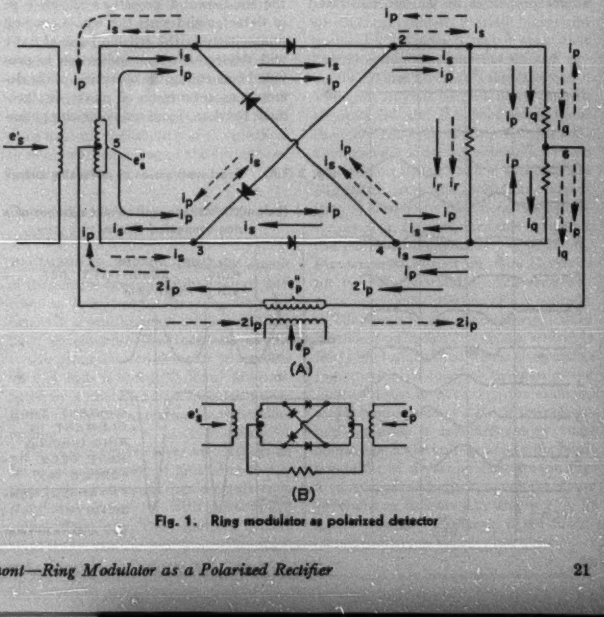
Synopsis: The performance of the ring modulator as a phase-sensitive rectifier is described qualitatively. Under the assumption of idealized conditions, and with the use of the analytical approach originally developed for the ring modulator, the polarized rectifier is described quantitatively, both for steady state and transient state. The analytical findings are checked experimentally. A differential polarized rectifier functioning as a true null detector is also discussed.

WHENEVER it becomes necessary to establish the absolute steady-state equality between two electrical phasors, a null-indicating device, or a null detector, must be used. Many such devices are in use to indicate the balance condition of bridges and, in a broader field, to permit the absolute measurement of phasor relationships.

Attention was focused early on the ring modulator, so successfully used in carrier communication, for, if instead of employing two different frequencies, i.e., the signal and carrier frequencies, the two are permitted to merge into a single measuring frequency, a new and powerful tool, the polarized rectifier, is obtained. Although the device has been previously described in the literature,[1,2] it is felt that it has been neither comprehensively treated nor given the recognition it deserves as one of great inherent usefulness. It is therefore proposed to cover this subject rather thoroughly in the following discussion.

Qualitative Aspects

The arrangement illustrated in Fig. 1(A) is diagrammatically identical to the ring modulator, although it differs functionally from the ring modulator arrangement. Fundamentally, four nonlinear resistance elements (1-2, 2-3, 3-4, 4-1), either of the vacuum-tube or the semiconductor type, are connected in series forming a "ring." (In the following, only the semiconductor type will be considered. All work was done with the Western Electric type-*D-98914* cuprous-oxide–copper varistor.) During the positive 1/2-cycle of signal e_s'' the resistance of one pair of elements is substantially reduced; during the negative 1/2-cycle, that of the other pair undergoes such a reduction, the medium producing these (exactly synchronous) reductions being an auxiliary potential, the polarizing potential e_p'', whose frequency is identically the same as that of the signal and whose amplitude is substantially in excess of that of the signal. In what manner rectification takes place will become apparent upon study of current flow for each 1/2-cycle, identified by solid and dashed arrows. Polarizing energy coupled into the circuit via an isolation transformer produces current $2i_p$ (solid), which splits at 5 into two equal components i_p. It will be assumed, for simplicity, that the resistance of an element in the blocking direction is infinite and that the characteristics of the elements are identical among each other. One component i_p thus greatly reduces the resistance of element 1-2, the other one that of 3-4: the pair of elements 1-2 and 3-4 are "unlocked" by the polarizing energy while the pair 2-3 and 4-1 becomes "forbidden" for energy transfer. Signal energy, likewise coupled by means of an isolation transformer, establishes current i_s (solid) freely flowing through 1-2, splitting into components i_q and i_r, recombining into i_s, and freely continuing through the unlocked element 4-3. For the next 1/2-cycle, polarizing current $2i_p$ (dashed) unlocks elements 2-3 and 4-1, forbidding passage through the pair 1-2 and 3-4. Signal energy, which has also undergone a synchronous phase reversal, establishes current i_s (dashed) flowing through 3-2, splitting into i_q and i_r and continuing through 4-1. It is at once realized that the direction of i_r and i_q remains invariant for either 1/2-cycle, and that hence the phenomenon of rectification takes place. A little thought will further reveal, that, if the polarity of e_p' is maintained while that of e_s' is reversed, the direction of i_r will likewise reverse: the phenomenon of passage through zero for phase reversal is guaranteed. Further study of the diagram indicates that for $e_s'=0$, no current flows through r (which represents the branch in which a d-c galvanometer



(A)

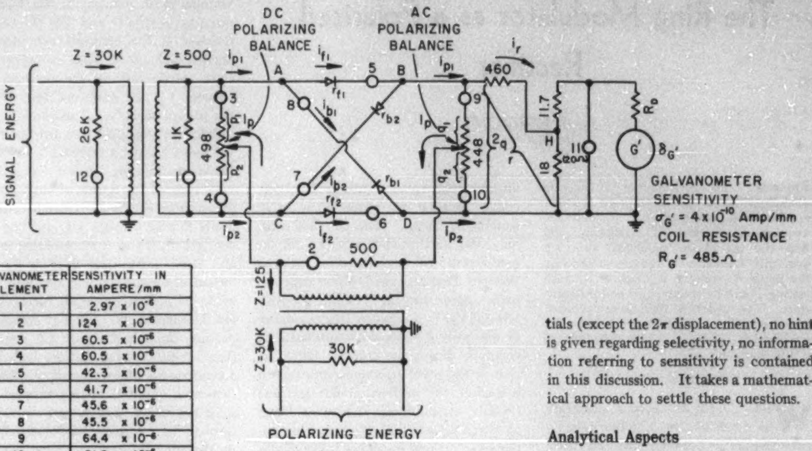
(B)

Fig. 1. Ring modulator as polarized detector

Paper 55-14, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted March 25, 1954; made available for printing October 29, 1954.

A. J. Hermont is with the Shell Development Company, Houston, Tex.

Gratitude is expressed to the Shell Development Company for permission to publish this paper.



GALVANOMETER ELEMENT	SENSITIVITY IN AMPERE/mm	
1	2.97×10^{-6}	
2	124×10^{-6}	
3	60.5×10^{-6}	
4	60.5×10^{-6}	
5	42.3×10^{-6}	
6	41.7×10^{-6}	
7	45.6×10^{-6}	
8	45.5×10^{-6}	
9	64.4×10^{-6}	
10	61.9×10^{-6}	
11	1.27×10^{-6}	(LOOKING INTO POINT H)
11	$.46 \times 10^{-6}$	(ELEMENT ALONE)
12	$.85 \times 10^{-6}$	

Fig. 2. Polarized rectifier for oscillography

is placed), since polarizing energy is exactly balanced out from that branch: the phenomenon of zero output for zero input also exists. The polarized rectifier indicated in Fig. 1(B) has the same desirable properties as the one illustrated in Fig. 1(A); however, sufficient elimination of polarizing energy from the galvanometer is not possible, especially if high sensitivity is demanded.

This qualitative discussion exhibits the fundamental properties of this type of detector and puts the whole a-c null proposition on the reliable basis of a d-c null device. But no indication is provided concerning the operation of the detector as a function of phase displacement between signal and polarizing potentials (except the 2π displacement), no hint is given regarding selectivity, no information referring to sensitivity is contained in this discussion. It takes a mathematical approach to settle these questions.

Analytical Aspects

The expression for the galvanometer deflection δ in terms of the harmonic signal strength E_{sm} and other physical factors has been derived in the Appendix. See equation 26.

The result may be summarized in terms of the following statements:

> For idealized semiconductor specifications, the d-c output current is strictly proportional to the real component of the signal voltage with respect to the polarizing voltage. If a voltage of different frequency is superimposed on the signal, it can contribute to the d-c output current only in so far as it is an exact odd harmonic of the signal frequency: in that case the additional contribution to the output is proportional to the real component of the signal voltage and inversely proportional to the order of the harmonic; the contributions attributable to the harmonics are either aiding or bucking the contribution attributable to the fundamental.
>
> All this reveals the paramount property that the output is not only sensitive to the real component of the signal phasor, but is also discriminative in favor of the signal frequency, so that unwanted energy is largely rejected. It is also seen very clearly that a null will obtain not only for the condition of zero input signal amplitude but also for a quadrature signal phase (with respect to the polarizing potential) and one must be on guard for this latter "fake" null.
>
> As to the a-c output, it has already been mentioned that the fundamental frequency is completely balanced out as far as polarizing energy is concerned. In the light of the analysis this statement becomes more precise. If, for simplicity, the uncommutated signal current is assumed to be $i_{R_{NC}} = I_{R_m} \cos m\omega t$, then the contribution to the a-c output current attributable to the nth harmonic of the commutating voltage is, by equation 7

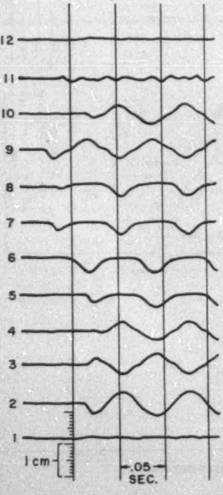
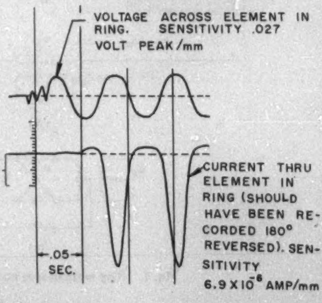


Fig. 3 (left). Exact suppression of polarizing energy

Fig. 4 (below). Voltage and current relations of a ring connected element



I VOLTAGE ACROSS ELEMENT IN RING. SENSITIVITY .027 VOLT PEAK/mm

CURRENT THRU ELEMENT IN RING (SHOULD HAVE BEEN RECORDED 180° REVERSED). SENSITIVITY 6.9×10^{-6} AMP/mm

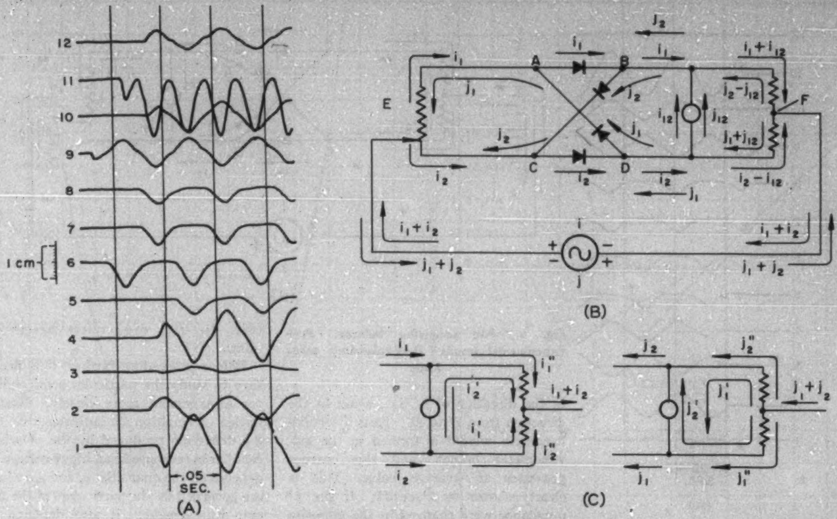


Fig. 5. D-c polarizing balance. A—Experimental results. B—Explanatory schematic. C—Formation of i_{12} and j_{12} in b

l. 74
1
55

$$i_{R_{mn_c}}=\frac{4}{\pi}\,\frac{I_{R_m}}{n}\cos m\omega t\cos n\omega t$$
$$=\frac{2}{\pi}\,\frac{I_{R_m}}{n}\{\cos(m-n)\omega t+\cos(m+n)\omega t\}\qquad(1)$$

If the signal frequency is exactly identical to the polarizing frequency, then $m=n=1$ and

$$i_{R_c}=\frac{2}{\pi}I_R(1+\cos 2\omega t)\qquad(2)$$

so that a second harmonic (as well as a d-c component) is expected. If a second harmonic is present in the signal, i.e., $m=2n$, a fundamental component, however, will result, in addition to a third-harmonic component as

$$i_{R_{mn_c}}=\frac{2}{\pi}I_{R_m}(\cos\omega t+\cos 3\omega t)\qquad(3)$$

These considerations are incorporated in the following statement: If the signal is distorted the output current contains harmonics and may contain the fundamental for the special case of a second harmonic distortion; moreover, as $m\rightarrow n$, slow beat frequencies will appear. It will be the function of the indicating device, i.e., the device capable of energy storage, to suppress these a-c components as much as possible.

The transient response is less favorable. According to equation 31 and the remark following it, every exponentially decaying step function is accompanied by a decaying train of square waves. If the polarizing frequency is low, the wave train will cause a bothersome fluctuation of the galvanometer and, unfortunately, no remedy for this defect exists.

Equation 26 represents an important result, the experimental verification of which it is proposed to study next. This will contribute to a full appreciation of the complexities involved and will shed some light on the validity of the idealizations of the functioning of the device used in the analysis.

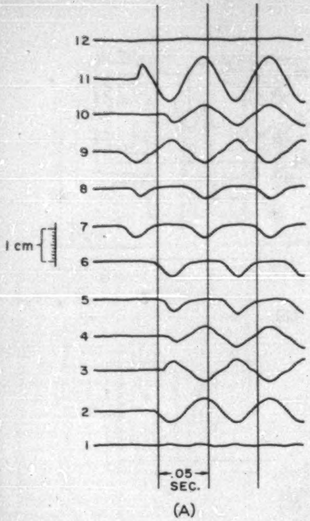
Experimental and Quantitative Aspects

Accordingly, Fig. 2, a special circuit of the general type under discussion, with suitable components for recording by means of a multichannel oscillograph has been developed. The locations of the twelve galvanometer elements (the resistance of each is about 20 ohms) are indicated by numerals 1 to 12, which correspond to the trace numbers on the oscillograms.

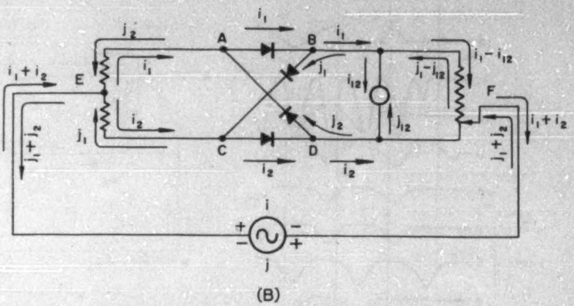
Fig. 3 is a tracing from an oscillogram for exact suppression of polarizing energy (frequency 14.3 cycles per second) from the output (trace 11) with zero-applied signal (trace 12). The polarizing voltage existing at the secondary of the polarizing transformer is 0.217-volt peak and is maintained throughout the tests. The d-c calibrations for the galvanometers are indicated in Fig. 2.

No fundamental is visible on trace 11 which does, however, indicate higher odd harmonics. Traces 5 and 6 show that flow of conduction currents through elements AB and CD occurs simultaneously, while conduction current flows through elements BC and DA simultaneously for the other 1/2-cycle of the polarizing voltage (traces 7 and 8 respectively). Phase reversals on traces 3 and 4, and on 9 and 10, result from the polarizing current flowing in opposite directions through resistors p_1 and p_2, and the same is true for the currents through resistors q_1 and q_2. Fundamental balance on trace 11 results for $p_1=248$ ohms and $q_1=229$ ohms, i.e., for a near-center position of the controls.

The reason for the appearance of a third harmonic on traces 3, 4, 9, 10, and 11 is fairly obvious. At all times the polarizing potential is equal to the sum of the voltage drops in p_1 and q_1, for example, and in semiconductor AB. If the polarizing potential is assumed purely sinusoidal and if the voltage across the semiconductor of the shape illustrated in Fig. 4 is subtracted from it, a voltage shape precisely as indicated by traces 3, 4, 9, and 10 results. Moreover, if the funda-



(A)



(B)

mental is balanced out, the shape indicated on trace 11 obtains.

Next, it is assumed that, still without feeding signal energy to the device, the p-balance is disturbed. Fig. 5(A) is an oscillogram for a drastic change: $p_1=498$ ohms, $p_2=0$. The result is not considered obvious, and it requires at least a qualitative analysis. A simplified network diagram is shown in Fig. 5(B), in which the polarizing energy for one 1/2-cycle feeds current i, and for the next 1/2-cycle, current j. The q-balance is shown at its center position, while the p-balance is upset as mentioned above. Thus, considering all ring elements to be identical, $i_2>i_1$, no matter how far the tap-point E is moved toward the bottom. For example, from Fig. 5(A), and with the use of the calibrations shown in Fig- 2, the maximum values of these currents are: $i_2=\delta_6\times\sigma_6=7.5\times41.7\times10^{-6}=313\times10^{-6}$ ampere and $i_1=\delta_5\times\sigma_5=5\times42.3\times10^{-6}=212\times10^{-6}$ ampere. For simplicity, it is assumed that branches AD and CB are totally blocking, so that currents i_2 and i_1 issue from points D and B respectively. Referring now, to Fig. 5(C), it is obvious that the current in the galvanometer branch becomes $i_{12}=i_2'-i_1'$ (the currents i_2 and i_1 flowing simultaneously) and, since $i_2'>i_1'$, the direction of i_{12} (Fig. 5(B)) is from D to B. For the j 1/2-cycle, the distribution of the currents is as shown, with $j_2>j_1$. Thus, the current in the galvanometer branch becomes $j_{12}=j_2'-j_1'$, in the direction of j_2', i.e., again in the direction from D to B. Thus a second-harmonic current is formed in the galvanometer branch, and this current possesses an average value. This is clearly shown on trace 11. If the p-unbalance were changed in the opposite direction, the polarity of the d-c component would reverse. These experiments indicate that the d-c component, along with the second-harmonic component, may be completely "zeroed" by an adjustment of the p-potentiometer: for then $i_2=i_1$ and $j_2=j_1$ and points B and D are at the same potential. This type of balance is loosely called the d-c polarizing balance (Fig. 2).

Fig. 6. A-c polarizing balance. A—Experimental results. B—Explanatory schematic

When next the q-balance is disturbed, no visible d-c component develops on trace 11, but fundamental frequency is evidenced, as seen from Fig. 6(A). The situation is illustrated by the schematic of Fig. 6(B), which shows the q-potentiometer drastically unbalanced (tap-off point F at bottom). Here $i_2>i_1$, and i_2 flows from D back to the generator; i_1 sends a component i_{12} through the galvanometer branch. For the j 1/2-cycle, j_2 flows directly to D while j_1 sends a component j_{12} in the opposite direction through the galvanometer branch; for identical elements $i_{12}=-j_{12}$. Cancellation, however, does not occur since these currents flow in time-quadrature; hence a near sine wave results. If the tap F is adjusted correctly, then $i_1=i_2$ and $j_1=j_2$, so that the voltage drop from F to B is equal to the one from F to D, and current in the galvanometer branch vanishes. This experiment teaches that the fundamental of the polarizing current may be completely zeroed by an adjustment of the q-potentiometer. This type of balance is referred to as the a-c polarizing balance (Fig. 2). Fig. 6(A) refers to $q_1=200$ ohms.

Before proceeding further, it is necessary to study the particular semiconductor arrangement more closely. Examination of equation 26 indicates that the d-c deflection produced by the polarized rectifier in response to an input voltage E_s depends on the quantities ρ_0 and g_G, which are governed by the particulars of the circuit arrangement; it also depends on other quantities irrelevant to the present discussion, and associated with the nature of the input signal itself and with the properties of the null-indicating device. But ρ_0 and g_G, by equations 16, 19, and 23 are functions of S and thus, in the last analysis, of the crucial quantities r_f and r_b by equation 17. Therefore, the question arises: what constant values should be assigned to these quantities? Of course, r_f and r_b are known to be nonlinear functions of time, but the analysis ignored this fact and assumed them to be constant. It must also be understood that the only information assumed to be available at the outset is the static current-voltage characteristic of the particular unilateral semiconductor under investigation, usually published by the manufacturer. For the Western Electric type-*D-98914* this relation is revealed for the forward 1/2-cycle and the backward 1/2-cycle by the "current" curves of Fig. 7(A), and (B) respectively. The static resistance being of no concern here, graphical differentiation of the current functions results in the dynamic-resistance functions, also shown in Fig. 7.

It should furthermore be assumed that the only information available regarding the voltage across either of the four semiconductor elements is its peak value, which is a design specification. That the voltage is not sinusoidal is illustrated by Fig. 4, which was obtained by using a d-c amplifier for the recording of the latter. It is somewhat surprising that the voltage

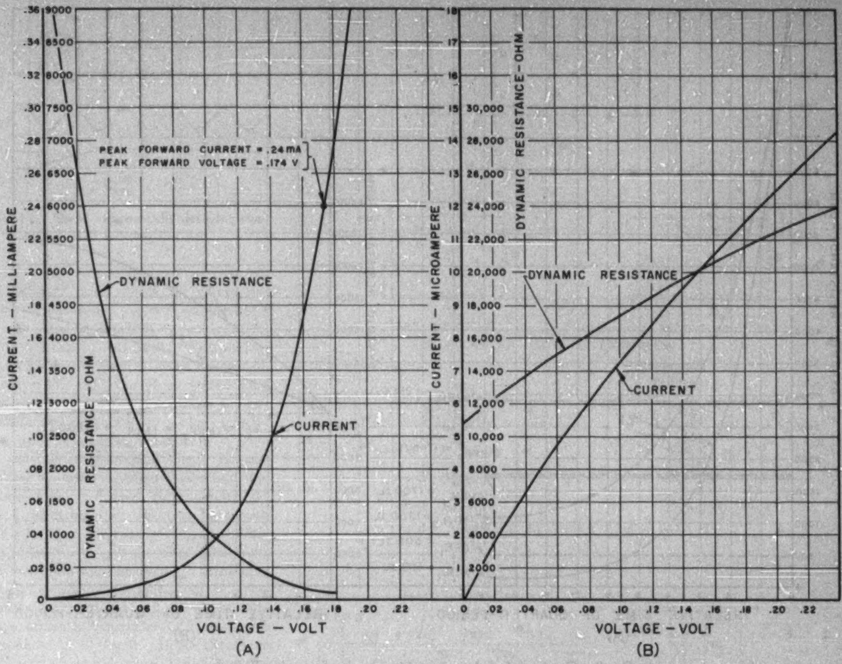


Fig. 7. Half-cycle semiconductor characteristics, Western Electric type D-98914. A—Forward. B—Backward

is so nearly sinusoidal. As an explanation of this phenomenon, it will be observed, by reference to Fig. 2, that the polarizing current I_p splits into two components i_{p_1} and i_{p_2}; i_{p_1} splits at A into i_{b_1} and i_{f_1}, while i_{p_2} splits at C into i_{b_2} and i_{f_2}, where the i_f-currents are understood to greatly exceed the i_b-currents. Assuming the potential at point D to be zero, the instantaneous potential of point A becomes $+i_{b_1}r_{b_1}$ if branch AD is considered; if branches AB and BC are considered, it becomes $+i_{f_1}r_{f_1}-i_{b_2}r_{b_2}+i_{f_2}r_{f_2}$. Thus it follows, by equating the potentials at A, that $i_{b_1}r_{b_1}+i_{b_2}r_{b_2}=i_{f_1}r_{f_1}+i_{f_2}r_{f_2}$ and, for identical elements, that

$$i_b r_b = i_f r_f \tag{4}$$

which implies that in this arrangement there exists a "constraint" on the instantaneous voltage across an element, forcing the voltage to remain invariant for either half-cycle. This is why the voltage-time distribution is so symmetrical with respect to the time axis. With this in mind and for purposes of further simplification of the problem, it is now assumed that the voltage-time distribution across an element is strictly sinusoidal.

To restate the immediate objective: It is required to find two quantities r_f and r_b, both constant over the respective half-cycle, such that when they are introduced into the proper equations, they will yield the d-c deflection produced by the polarized rectifier as demanded by equation 26. But this is manifestly contrary to the physics of the situation, since it has been shown that the dynamic resistances are by no means constant as functions of voltage and, hence, as functions of time. The simplifying assumptions made in the analysis are nevertheless justifiable, as will be subsequently shown. These assumptions, of course, were made in order to avoid the difficulties of the nonlinear mathematics involved in the problem. It should be stressed that resort to empiricism is deliberately undertaken at this point and that all claims to rigor are relinquished as is the expectancy of perfect agreement between analysis and experiment.

Returning to the conditions of the experiment, it is seen from the calibrated Fig. 4 that the peak forward current is 0.24 milliampere (ma) and the peak forward-and-backward voltage about 0.174 volt. The distribution of the forward-and-backward dynamic resistances, plotted against a relative quarter-period time scale and based on the foregoing sinusoidal voltage-time distribution, is revealed by curves 1 of Fig. 8. The minimum forward resistance is 150 ohms, the minimum backward resistance 10,800 ohms. The question is, what effective forward-and-backward resistance may be deduced from this information?

For the forward quarter-period it is felt that the minimum resistance should play a part; also that the time-average resistance should be involved. The empirical formulation is now made that the effective forward resistance shall be the geometric mean between the time-average and the minimum resistances. Accord-

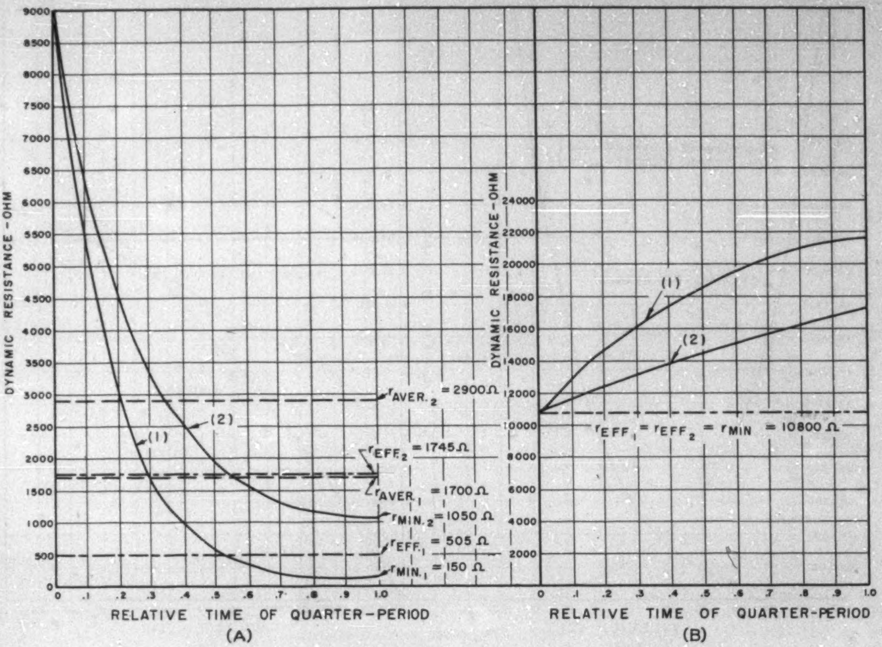


Fig. 8. Quarter-period dynamic resistance. A—Forward. B—Backward

ingly, since graphical integration of curve 1, Fig. 8(A) results in an average resistance of 1,700 ohms, the effective forward resistance becomes $r_f=505$ ohms. For the backward quarter-period, it is felt that the rise of dynamic resistance over its value at the transition between the forward-and-backward quarter-periods should be ignored; this, too, is based on empirical procedures. Accordingly, the effective backward resistance becomes $r_b=10,800$ ohms. With these values decided upon, the computation of δ from equation 26 proceeds without difficulty.

Referring to Fig. 9(A) i.e., to the case where the signal voltage is the fundamental and the phase shift is zero (or 180 degrees), equation 26 becomes $\delta=(2/\pi)(1/\sigma_G)g_G\rho_0E_s$. For the transformer indicated in Fig. 2, the experimentally determined constants are $L_1=387$ henrys, $R_1=1,800$ ohms, $R_2=52$ ohms, $\nu=7.75$. By equation 17 $S=2,335$. The fourpole resistance R_P may be computed from equation 19 if $R=2qr//2q+r=448\times 471.5/919.5=229.7$ ohms is substituted, with the result $R_P=1,145$ ohms. Thus the total load resistance R_o becomes 257.5 ohms (parallel combination of R_P, 1,000 ohms, 498 ohms).

Therefore, by equation 16 $\rho_0=9.82\times 10^{-2}$, when the frequency 13.4 cycles is given consideration. Equation 23 yields $g_G=3.72\times10^{-4}$ mho, while E_s is taken from Fig. 9(A) from the deflection of the calibrated galvanometer 12, namely $E_s=\delta_{12}\sigma_{12}26\times10^3=7\times0.85\times10^{-6}\times26\times10^3=$

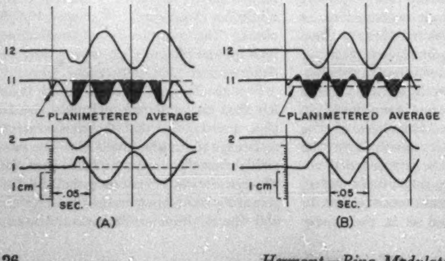


Fig. 9 (left). Signal voltage. A—In phase with polarizing voltage. B—Lagging behind polarizing voltage

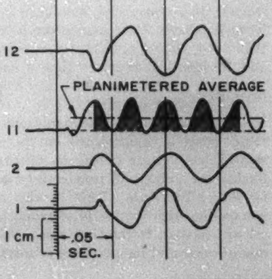


Fig. 10 (right). Signal voltage containing third harmonic distortion

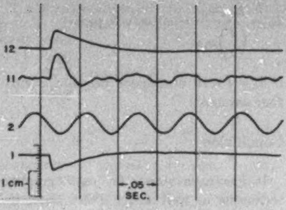
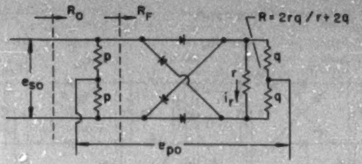
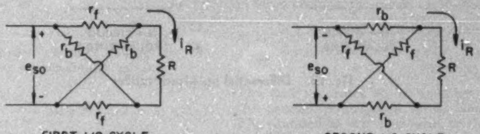


Fig. 11. Signal exponentially decaying step function

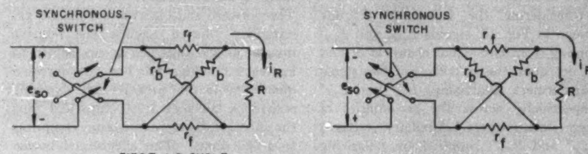


(a) POLARIZED RECTIFIER



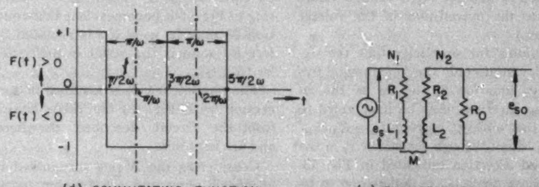
FIRST 1/2 CYCLE SECOND 1/2 CYCLE

(b) POLARIZED RECTIFIER EQUIVALENT I



FIRST 1/2 CYCLE SECOND 1/2 CYCLE

(c) POLARIZED RECTIFIER EQUIVALENT II



(d) COMMUTATING FUNCTION (e) TRANSFORMER COUPLING

Fig. 12. Polarized rectifier analysis

0.1547-volt peak. With the galvanometer sensitivity $\sigma_G=\sigma_{11}=1.27\times10^{-6}$ finally, analysis yields $\delta_A=2/\pi\times1/1.27\times10^{-6}\times3.72\times10^{-4}\times9.82\times10^{-2}\times0.1547=$ 2.84 millimeters (mm). Next, the shaded area in Fig. 9(A) was planimetered and, as an average of seven readings, the experimental deflection by means of oscillography becomes $\delta_{EO}=4.09$ mm.

Another experimental check is provided by observing the d-c deflection of a long-period sensitive galvanometer G' and adjusting the deflection by means of a decade resistor R_D (see Fig. 2). With $R_D=5{,}000$, a galvanometer deflection $\delta_{G'}=16$ mm was observed. Taking into account the coil resistance and the sensitivity indicated in Fig. 2, the direct voltage across galvanometer 11 is computed to be 3.51×10^{-6} volt. Hence the direct current through element 11 is 1.755×10^{-8} amperes and, taking into account the sensitivity of element 11, the experimentally obtained deflection by means of a long-period galvanometer becomes $\delta_{EG'}=3.815$ mm.

Fig. 9(B) illustrates a case where phase shift is involved. Directly from the oscillogram the signal voltage (trace 1) lags behind the polarizing voltage (trace 2) by 54 degrees, if trace 1 is considered recorded 180 degrees out of phase. The oscillogram trace 12 furnishes $E_s=0.1492$-volt peak; hence $E_s\cos\varphi=0.0877$-volt peak. The other quantities remaining as in the previous case, analysis furnishes $\delta_A=1.61$ mm. The oscillographic approach yields $\delta_{EO}=2.10$ mm, while the long-period galvanometer provides $\delta_{EG'}=$ 1.91 mm.

If a strong third harmonic is present in the signal, a situation revealed by Fig. 10 results. Observe that since the signal phase (trace 1) is reversed as compared to the signal in Fig. 9(A), the d-c component on trace 11 is also reversed as expected. A Fourier analysis of the signal on trace 12 was made, the results being for $m=1$, 3,5: $E_{s1}=0.148$, $\varphi_1=-9.25$ degrees; $E_{s3}=0.0151$, $\varphi_3=+77.38$ degrees; $E_{s5}=$ 0.00638, $\varphi_5=-80.37$ degrees. If for the higher harmonics the asymptotic value of $\rho_{om}=0.0983$ is used

$$\sum_{m=1}^{m\,\mathrm{odd}}\frac{E_{sm}\rho_{om}}{m}\cos\varphi_m=0.148\times9.82\times$$

$$10^{-2}\times0.9871-\frac{1}{3}\times0.0151\times9.83\times$$

$$10^{-2}\times0.2181+\frac{1}{5}\times0.00638\times9.83\times$$

$$10^{-2}\times0.1668=+0.014263$$

so that $\delta_A=2.66$ mm. The oscillogram yields $\delta_{EO}=3.84$ mm and the long period galvanometer $\delta_{EG'}=3.82$ mm.

If the results observed by means of the long-period galvanometer are considered correct, it must be concluded that the analytic results are at the most about 30 per cent lower than the true ones, i.e., they are on the conservative side.

Fig. 11 illustrates the response to an exponentially decaying step function, which is seen to be in complete accord with equation 31.

All these results are very gratifying, for they prove that the operation of the device is clearly understood and that the idealizing assumptions are justified in view of the close agreement between theory and experiment.

Principles of the Differential Polarized Rectifier

It is possible to interconnect two polarized rectifiers of the type shown in Figs. 1 and 12 in such a manner that if each is fed by a signal and a corresponding auxiliary signal (polarizing voltage), the detecting device (d-c galvanometer) will

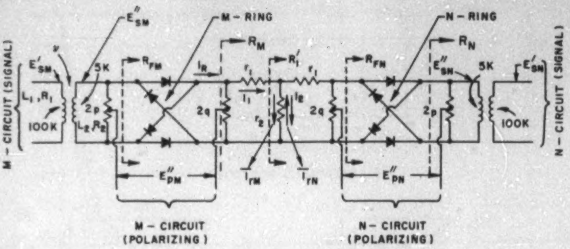


Fig. 13. Differential polarized rectifier

indicate zero current under conditions to be described.

Considering the schematic diagram shown in Fig. 13, signals E_{sM}' and E_{sN}' coming from two signal channels *M* and *N*, act upon the primaries of two signal transformers, establishing E_{sM}'' and E_{sN}'' respectively, across the secondaries of these transformers. Auxiliary voltages E_{pM}'' and E_{pN}'', supplied by polarizing channels *M* and *N*, are established at the (adjustable) mid-points of the respective resistors $2p$ and $2q$. (All voltage values refer to the magnitudes of the voltage phasors.)

Assuming for simplicity that the signals E_s contain only the fundamental frequency, equation 25, as far as the *m*-side is concerned, may be interpreted to mean that a direct current $I_{rM}=K_M E_{sM}' \cos \varphi_M$ flows through resistance r_2 in the assumed direction indicated in Fig. 13. The phase angle φ_M is understood to be the one between the phasors e_{sM}'' and e_{pM}''. Assuming now that the polarity of the voltages applied to the *N*-side is the same as that of the corresponding voltages of the *M*-side and that the ring configurations are connected as indicated, it becomes obvious that $I_{rN}=K_N E_{sM}' \cos \varphi_N$ will flow through r_2 in the opposite direction. Thus $I_r=I_{rM}-I_{rN}=0$ only if

$$K_M=K_N$$

and

$$E_{sM}' \cos \varphi_M=E_{sN}' \cos \varphi_N$$

That such a differential polarized rectifier may be employed as a basis for a true null indicator is described elsewhere.[3] It will now be shown numerically how the analysis developed in the Appendix can be utilized to determine the minimum value E_s' (say E_{sM}') required under special conditions.

Suppose the peak voltage across one of the elements is 0.1 volt. Then the peak forward current, according to Fig. 7(A) is only 0.032 milliampere and the minimum dynamic resistance is 1,050 ohms. The forward-and-backward dynamic resistances, plotted against the relative quarter-period time scales, are shown as curves 2 in Fig. 8. For the forward quarter-cycle the time-average dynamic resistance becomes 2,900 ohms and thus the effective dynamic resistance computes to 1,745 ohms. The effective dynamic resistance for the backward quarter-cycle is, as previously, 10,800 ohms. Hence, by equation 17, $S=4{,}350$ ohms. Referring to Fig. 13 it becomes clear that equation 23 may be used if r is replaced by (r_1+R_1') and if the result is multiplied by (R_1'/r_2).

The numerical values assumed in connection with Fig. 13 are those arising from the circuit described elsewhere[3] and are kept here.

Considering the *N*-ring terminated in $2p=5{,}000$ ohms and a 5,000-ohm transformer impedance in parallel, i.e., in $R_N=2{,}500$ ohms equation 19 yields $R_{FN}=3{,}940$ ohms. Thus, with $2q=5{,}500$ ohms, $r_1=5{,}000$ ohms and $r_2=20{,}500$ ohms, $R_1'=5{,}380$ ohms is obtained. Hence, equation 23 should be multiplied by $(5{,}380/20{,}500)=0.2625$, and $r=10{,}380$ should be used. Thereupon $g_G=0.994\times 10^{-6}$ mho. For the calculation of ρ_0 from equation 16, the *M*-ring is seen to be terminated in $R_M=3{,}590$ ohms, so that $R_{FM}=4{,}200$ ohms and $R_0=2{,}280$ ohms. If the frequency is taken as 10 cycles per second and the experimentally obtained transformer constants are $L_1=19{,}510$ henrys, $R_1=31{,}000$ ohms, $R_2=2{,}150$ ohms, and $\nu=4$, there results $\rho_0=0.08975$.

Next, it will be assumed that detection of ±1/2 mm is possible. Then, to detect variations with an error of 1 per cent, the deflection should be $\delta_{\min}=50$ mm. Therefore equation 26 yields the minimum required voltage E_s' at the primary of the transformer

$$E_s'=\frac{\delta_{\min}\,\sigma_G}{\left(\frac{2}{\pi}\right)g_G\rho_0}=0.0346 \text{ volt peak}$$

if $\sigma_G=4\times 10^{-10}$ ampere per mm is further assumed.

Conclusion

It has been shown by oscillographic technique as well as by direct observation of a long-period d-c galvanometer that the analytical expression derived for the d-c response of a polarized rectifier (involving errors sufficiently small) represents a reliable result. This was demonstrated for various types of signals. However, it is necessary to follow some empirical rules in numerical work, unless an unwarranted complexity involving nonlinear mathematics is introduced. The analytically obtained transient response is also corroborated by experiment. It is therefore felt that this discussion adds materially to a better understanding of the mechanism of operation of the polarized rectifier of the ring modulator type.

Vol.
Pt. 1
1955

Appendix. Communication, Galvanometer, and Other Physical Factors

In the polarizing rectifier, Fig. 12(A), the polarizing circuit may be omitted entirely if only its effect on the ring is preserved. In the equivalent circuits *I*, Fig. 12(B), the semiconductor lattice is replaced by a pair of relatively low resistors r_f (forward or conducting) and a pair of relatively high resistors r_b (backward or blocking), assumed to be constant during the entire 1/2-cycle and changing location (within the lattice) abruptly as the signal voltage e_{s0} changes polarity between 1/2-cycles. For either 1/2-cycle, signal current flows mainly through r_f resulting in the invariance of the direction of i_R. The polarizing voltage e_{p0} in conjunction with the ring thus acts exactly as if a double-pole double-throw switch were operated synchronously, commutating the input voltage e_{s0} but leaving the lattice intact: this is illustrated by the equivalent circuit *II*, Fig. 12(C).

This point of view leads to the mathematical formulation that it is only necessary to multiply the output current by a time function $F(t)$, the commutating function, having the shape shown in Fig. 12(D), to obtain the combined effect of the interaction of the polarizing voltage and the semiconductor ring circuit upon the signal voltage. $F(t)$ has the value +1 for one 1/2-period of duration π/ω and the value −1 for the next 1/2-period. This concept was first applied to modulator theory[4] and is here adapted to the polarized rectifier, which is merely a special case of the latter.

The Fourier series for the function $F(t)$ may be obtained without difficulty, the result being

$$F(t)=\frac{4}{\pi}\sum_{n=1}^{n\,\mathrm{odd}}\pm\frac{\cos n\omega t}{n}=\frac{4}{\pi}\left(\cos\omega t-\frac{1}{3}\cos 3\omega t+\frac{1}{5}\cos 5\omega t\ldots\right)\qquad(5)$$

Without loss of generality, the signal output current when no commutation takes place may be assumed to be

$$i_{R_{NC}}=\sum_{m=1}^{m}I_{Rm}\cos(m\omega t+\varphi_m)\qquad(6)$$

Now, to obtain the effect produced by commutation, it is necessary only to perform the operation

$$i_{RC}=i_{R_{NC}}F(t)\qquad(7)$$

for each component of the two series.

The formulation for the rectified component of i_{RC} is given in terms of an average for an infinitely long time, i.e.

$$I_{R_{DC}}=\lim_{T\to\infty}\frac{1}{T}\int_0^T i_{RC}(t)dt\qquad(8)$$

The contribution of the nth harmonic of $F(t)$ and the mth harmonic of $i_{R_{NC}}$ upon $I_{R_{DC}}$ is therefore

$$I_{R_{nm_{DC}}}=\lim_{T\to\infty}\frac{1}{T}\frac{4}{\pi}\frac{I_{Rm}}{n}\int_0^T\pm\cos n\omega t\cos\times(m\omega t+\varphi_m)dt\qquad(9)$$

The definite integral is

$$J=\int_0^T\cos n\omega t(\cos m\omega t\cos\varphi_m-\sin m\omega t\sin\varphi_m)dt=\frac{1}{2}\cos\varphi_m\times\left(\frac{\sin(m-n)\omega t}{\omega(m-n)}+\frac{\sin(m+n)\omega t}{\omega(m+n)}\right)_0^T+\frac{1}{2}\sin\varphi_m\left(\frac{\cos(m-n)\omega t}{\omega(m-n)}+\frac{\cos(m+n)\omega t}{\omega(m+n)}\right)_0^T\qquad(10)$$

Since the sin- and cos-functions have a bound at ±1, examination of equation 10 reveals that no matter how closely $m\to n$. but as long as $m\neq n$, J remains $<\infty$ and hence, in the limit, equation 9 yields $I_{R_{nm_{DC}}}=0$. To obtain the limit for $m=n$ it is best to substitute first $\epsilon=m-n$ into equation 10 and then to go over to the limit $\epsilon\to 0$

$$J=\lim_{\epsilon\to 0}\left[\frac{1}{2}\cos\varphi_m\left(\frac{\sin\epsilon\omega T}{\omega\epsilon}+\frac{\sin(\epsilon+2n)\omega T}{\omega(\epsilon+2n)}\right)+\frac{1}{2}\sin\varphi_m\left(\frac{\cos\epsilon\omega T}{\omega\epsilon}+\frac{\cos(\epsilon+2n)\omega T}{\omega(\epsilon+2n)}-\frac{1}{\omega\epsilon}-\frac{1}{\omega(\epsilon+2n)}\right)\right]$$
$$=\frac{1}{2}\cos\varphi_m T+\left(\frac{1}{2}\cos\varphi_m\frac{\sin 2n\omega T}{2n\omega}+\frac{1}{2}\sin\varphi_m\frac{(\cos 2n\omega T-1)}{2n\omega}\right)\qquad(11)$$

If this is substituted into equation 9, the limiting process $T\to\infty$ is completed and it is remembered that m merges with n, there obtains

$$I_{Rm_{DC}}=\pm\frac{2}{\pi}\frac{I_{Rm}}{m}\cos\varphi_m\qquad(12)$$

Therefore, the total contribution to the rectified current is

$$I_{R_{DC}}=\frac{2}{\pi}\sum_{m=1}^{m\,\mathrm{odd}}\frac{I_{Rm}}{m}\cos\varphi_m,\ m=1,-3,+5,-7,\ldots\qquad(13)$$

and the direct current flowing in the galvanometer branch finally becomes

$$I_{r_{DC}}=\frac{2}{\pi}\sum_{m=1}^{m\,\mathrm{odd}}\frac{I_{rm}}{m}\cos\varphi_m,\ m=1,-3,+5,-7,\ldots\qquad(14)$$

Referring now to the transformer coupling of the signal energy, Fig. 12(E), it will be assumed that the magnetic paths which the flux must take in linking the primary, the secondary, and the mutual turns are identical: this implies zero leakage flux,[8] an idealized condition assumed to exist for the frequency range under consideration and, while never met in practice, nevertheless resulting in a fair approximation; from the standpoint of analysis the simplification is, of course, enormous. Thus, letting

$$\begin{aligned}L_1&=K_1N_1^2\\L_2&=K_2N_2^2\\M&=KN_1N_2\end{aligned}\qquad(15A)$$

where K is essentially the permeances of the flux path and N the primary and secondary turns, and stipulating that

$$K_1=K_2=K\qquad(15B)$$

there results

$$L_2=L_1\frac{1}{\nu^2}\qquad(15C)$$

and

$$M=\sqrt{L_1L_2}=L_1\frac{1}{\nu}\qquad(15D)$$

where ν is the primary-to-secondary turns ratio. For such a transformer the magnitude of the output voltage E_{so} in terms of the input voltage E_s takes on the following expression, easily derivable from the coupled circuit of Fig. 12(E)

$$E_{so}=E_s\times\frac{\nu\omega L_1R_o}{([R_1(R_2+R_o)\nu^2]^2+(\omega L_1)^2[R_1+(R_2+R_o)\nu^2]^2)^{\frac{1}{2}}}=E_s\rho_o\qquad(16)$$

Here ρ_o is the magnitude of the transmission ratio and R_o = total load resistance upon which the transformer works [see Fig. 12(A].

The resistance looking into the terminated four-pole R_ν of Fig. 12(A), is available in terms of the image characteristic resistance S and the image transfer factor, ζ from general network theory.[8] Both of these are real quantities in this particular case and are expressible in terms of the lattice elements r_f and r_b. Steady-state lattice theory is convenient and justifiable because of the simplifying assumptions of constant, nonreactive lattice elements; transient state, as far as the loaded four-pole is concerned, thus does not exist as the ring elements are commutated. For the symmetrical lattice under consideration

$$S=(r_fr_b)^{\frac{1}{2}}\qquad(17)$$

$$\zeta=\ln\left(\frac{(r_b/r_f)^{\frac{1}{2}}+1}{(r_b/r_f)^{\frac{1}{2}}-1}\right);\ e^{-\zeta}=\frac{S-r_f}{S+r_f}\qquad(18)$$

In terms of these, the resistance R_ν becomes

$$R_\nu=S\frac{R+S\tanh\zeta}{S+R\tanh\zeta}=\frac{R(S^2+r_f^2)+2S^2r_f}{(S^2+r_f^2)+2Rr_f}\qquad(19)$$

The load resistance R_o in equation 16 is the parallel combination of $2p$ and R_ν. The situation is now equivalent to the one in which a generator with electromotive force $|e_{so}|$ and with zero internal impedance feeds the lattice. For this case the magnitude of the output current, I_R, in terms of S and ζ becomes

$$I_R=E_{so}\frac{2}{S+R}e^{-\zeta}\frac{1}{1-\dfrac{S-R}{S+R}e^{-2\zeta}}\qquad(20)$$

If equation 18 is substituted for $e^{-\zeta}$ and $e^{-2\zeta}$, equation 20 yields

$$I_R=E_{so}\frac{S^2-r_f^2}{R(S^2+r_f^2)+2S_{r_f}{}^2}\qquad(21)$$

Since, furthermore

$$I_r=I_R\frac{R}{r}=I_R\frac{2q}{r+2q}\qquad(22)$$

the result becomes

$$I_r=E_{so}\frac{(S^2-r_f^2)}{r(S^2+r_f^2)+S^2r_f\left(\dfrac{r+2q}{q}\right)}=E_{so}g_G\qquad(23)$$

where g_G = transfer conductance, transferring a voltage applied to the input of the lattice to the galvanometer current.

Combining equations 16 and 23 and considering the mth harmonic

$$I_{rm}=E_{sm}\rho_{om}g_G\qquad(24)$$

Substituting this into equation 14, the d-c galvanometer current is brought in the form

$$I_{r_{DC}}=\frac{2}{\pi}g_G\sum_{m=1}^{m\,\mathrm{odd}}\frac{E_{sm}\rho_{om}}{m}\cos\varphi_m,\ m=1,-3,+5,-7,\ldots\qquad(25)$$

It should be kept in mind that φ_m is understood to refer to the phase displacement existing between signal voltage e_{so} and polarizing voltage e_{po} at the secondaries of the isolation transformers, while E_{sm} refers to the signal amplitude at the primary of the signal isolation transformer.

Introducing the galvanometer sensitivity σ_G in amperes per mm, the expression for the galvanometer deflection δ in mm becomes

$$\delta=\frac{2}{\pi}\frac{1}{\sigma_G}\sum_{m=1}^{m\,\mathrm{odd}}\frac{E_{sm}\rho_{om}}{m}\cos\varphi_m,\ m=1,-3,+5,-7,\ldots\qquad(26)$$

Turning now to the transient response, it will be assumed that

$$i_{R_{NC}}(t) = I_R \epsilon^{-\alpha t} 1 \qquad (27)$$

i.e., that it represents an exponentially decaying unit function. Considering only one term of the summation of equation 25, equation 7 becomes

$$i_{R_C}(t) = \frac{4}{\pi} I_R \frac{1}{n} \epsilon^{-\alpha t} \cos n\omega t \, 1 \qquad (28)$$

Using the appropriate identity of operational calculus transforming a time function into an equivalent operator

$$\epsilon^{-\alpha t} \cos n\omega t \, 1 = \frac{p(p+\alpha)}{(p+\alpha)^2 + n^2\omega^2} 1 \qquad (29)$$

the operational response then becomes the following

$$i_{R_C}(p) = \frac{4}{\pi} I_R \frac{\dfrac{p(p+\alpha)}{n}}{(p+\alpha)^2 + n^2\omega^2} 1 \qquad (30)$$

the solution of which, by means of the expansion theorem, is

$$i_{R_C}(t) = \frac{4}{\pi} I_R \frac{1}{n} \epsilon^{-\alpha t} \cos n\omega t \qquad (31)$$

which means that the response to an exponentially decaying step function is an exponentially decaying cosine wave. However, the commutating function, equation 5, is a summation of cosine waves producing a square wave. Thus the complete transient response will be an exponentially decaying square wave.

References

1. Electronic Null Detectors for Use with Impedance Bridges, Horatio W. Lamson. *AIEE Transactions*, vol. 66, 1947, pp. 535–40.

2. Phase Indicating Null Indicator for Bridges. *Electronics*, New York, N. Y., vol. 17, no. 8, Aug. 1944, pp. 242–60.

3. Precision Measurement of Phasors in Electrical Networks as a Function of Frequency, A. J. Hermont. *AIEE Transactions*, vol. 73, pt. I, Sept. 1954, pp. 365–75

4. Copper Oxide Modulators in Carrier Telephone Systems, R. S. Caruntheis. *Bell System Technical Journal*, New York, N. Y., vol. 18, 1939, pp. 315–37.

5. Communication Networks (book), E. A. Guillemin. John Wiley & Sons, Inc., New York, N. Y., vol. I, 1931.

6. Communication Networks (book), E. A. Guillemin. John Wiley & Sons, Inc., New York, N. Y., vol. II, 1935.

A New Telegraph Serviceboard Using Electronic Circuits

J. R. DAVEY
NONMEMBER AIEE

F. H. HANLEY
NONMEMBER AIEE

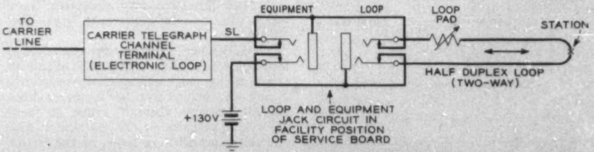
M. R. PURVIS
ASSOCIATE MEMBER AIEE

THE large increase in the use of teletypewriters in the Bell System in recent years has necessitated the development of improved arrangements for interconnecting telegraph facilities to form the networks employed by private-line services (also called leased-wire services). Greater flexibility is required in order to permit these networks to be built up and rearranged quickly to meet ever-changing customer requirements. With continuous improvements in the operating stability and reliability of telegraph line facilities, together with greatly increased demands from government and industry, private-line networks have become vastly complex.

Bell System telegraph service now uses the teletypewriter almost entirely, and manual Morse service is pratically negligible. Many networks extend over large areas of the nation, terminating in a number of cities. One employs about 85,000 miles of line facilities and another includes approximately 1,500 teletypewriter stations. These large networks are arranged for half-duplex service, which permits only one station to send at a time to all other stations on the network, and a large proportion of the stations are arranged for receiving-only operation. Such networks commonly serve government agencies, airlines, news agencies, financial institutions, and industrial organizations. Many operate at the higher teletypewriter transmitting speeds. Any false pulse or "hit" occurring on such a circuit will often interfere with transmission on the entire network unless it occurs on a 1-way branch, in which case its effect will be less serious. A disturbance of longer duration such as a steady space signal will interrupt service until located and removed. Even short interruptions to service on the more complex networks are likely to prove costly and result in serious inconvenience to the customer. Effort is necessary to reduce such occurrences to a minimum, and much development work has been directed toward improving present telegraph administrative arrangements which are performed largely at telegraph testboards. The results of this effort have culminated in the development of a new form of testboard known as a telegraph serviceboard which employs an electronic type of hub circuit to facilitate the interconnection and rearrangement of lines and loops for private-line telegraph service. (This is a multiple rather than the conventional series form of telegraph interconnection.)

Operating experience has indicated the importance of concentrating the telegraph administrative work in the hands of a minimum number of competent service attendants. This not only assures high efficiency in the performance of the work and faster restoration of service in case of trouble, but also provides economies in telegraph operating costs.

The new serviceboard accomplishes this centralization of effort by concentrating a larger work load in one position where it is easily accessible to an attendant. It employs a switchboard type of section with small strip-mounted jacks which enable an attendant to carry on his work from a seated position. The working load now handled at a single serviceboard position would occupy three or four positions of telegraph testboard. By



TO CARRIER LINE
CARRIER TELEGRAPH CHANNEL TERMINAL (ELECTRONIC LOOP)
SL
EQUIPMENT
LOOP
LOOP PAD
STATION
HALF DUPLEX LOOP (TWO-WAY)
+130V
LOOP AND EQUIPMENT JACK CIRCUIT IN FACILITY POSITION OF SERVICE BOARD

Fig. 1. Direct-leg termination of electronic carrier telegraph channel terminal in half-duplex loop

---

Paper 55-38, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31-February 4, 1955. Manuscript submitted October 21, 1954; made available for printing November 17, 1954.

J. R. Davey and M. R. Purvis are with the Bell Telephone Laboratories, Inc., New York, N. Y., and F. H. Hanley is with the American Telephone & Telegraph Company, New York, N. Y.

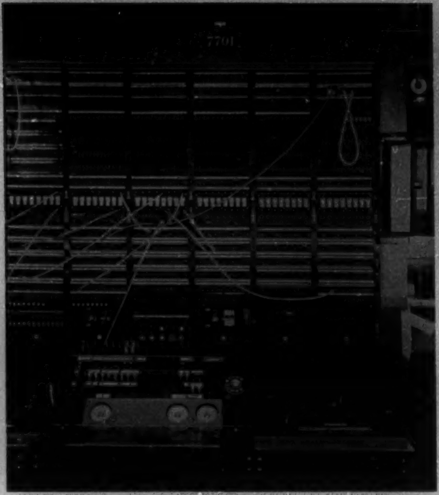


Fig. 2 (left). Facility position

Fig. 3 (above). Five service positions

ol. 74
. 1
55

utilizing the hub circuit principle and other advanced operating features, the work of the attendant is made simpler and easier and the objective of concentrating a higher work load in one position has been attained.

Service Features

GENERAL

The work at the serviceboard is mainly of two types; namely, that of establishing and rearranging private-line networks to meet customer requirements and that of testing and maintaining these networks and their component line and loop facilities. Line facilities or telegraph intercity trunks include a repeater at each terminal. They are usually capable of operating full duplex (transmitting in both directions simultaneously) but are most often employed for half-duplex service (transmitting in only one direction at a time). Loop facilities interconnect the serviceboard and the customer's station and include an electronic repeater at the serviceboard office for hub circuit operation.

Line facilities are often used interchangeably for private line or for trunks used in the teletypewriter exchange service (TWX) which is also furnished by the Bell System. In TWX service the facility interconnecting the customer's station and the TWX switching center is called a subscriber line circuit. TWX subscriber line circuits do not terminate in the telegraph serviceboard. Although the switching of TWX subscriber lines and trunk circuits is performed at TWX switchboards or by TWX automatic equipment, the assignment and interconnection of line facilities to TWX intertoll trunk circuit equipment and the maintenance of these facilities is a function of the serviceboard.

DIRECT-LEG VERSUS HUB CIRCUIT OPERATION

Two different circuit arrangements are employed for interconnecting line and loop facilities at a serviceboard. Networks which involve only two legs or branches, such as a line and a loop or two lines, are likely to require little service attention and are usually interconnected directly on a simple balanced loop or neutral basis using relay-type repeaters. This arrangement, which provides so-called "direct-leg" operation, is employed for all TWX intertoll trunk circuits in addition to 2-leg private-line networks. Networks involving three or more legs and all those requiring signal regeneration are interconnected by means of an electronic hub circuit in order to simplify network changes, to expedite testing and maintenance, and to permit the use of electronic regeneration.[1]

The direct-leg principle of interconnecting legs is essentially the same as the older conventional arrangements heretofore employed at telegraph testboards. The loop conductors are connected directly to the legs of the line repeater as shown in Fig. 1 and transmission is on a polar or neutral basis. Direct-leg connections require no electronic hub terminating equipment and as approximately 60 per cent of the networks (including TWX trunk circuits) include only two legs, the equipment savings are more than ample to justify the diversity of operating methods which are imposed by the two types of interconnection. Circuits connected on a direct-leg basis are terminated in facility positions of the serviceboard, whereas those arranged for hub circuit operation appear in service positions.

Facility and service positions of the telegraph serviceboard are shown respectively in Figs. 2 and 3. Fig. 4 shows a hub circuit network which includes two lines and a loop with optional provision for electronic regeneration. Each line facility terminates in an electronic hub adapter called a coupling unit, and the loop terminates in an equivalent adapter called a loop repeater. Fig. 5 shows these electronic units. While only three legs are shown in Fig. 4, such a hub may include a maximum of from 12 to 20 legs depending on certain factors which are discussed later. This form of operation is also half duplex, and when regeneration is provided, a single unidirectional regenerative repeater serves the hub, regenerating the signals from any leg to all other legs. This provides so-called multiway regeneration, which is more economical in the use of regenerative repeaters than older, conventional telegraph circuits.

SERVICEBOARD TERMINATING AND PATCHING ARRANGEMENTS

Telegraph line and loop facilities are normally cross-connected to form the network layout specified by the customer. When orders are received for new networks or modifications of existing layouts, the work is usually done quickly on a temporary basis by manual patching

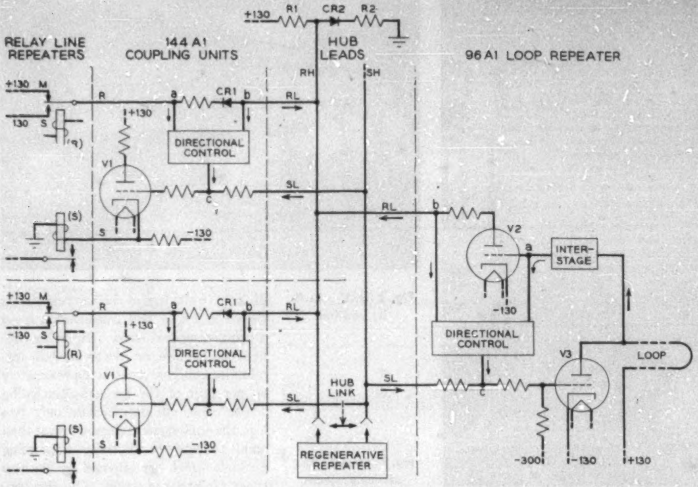


Fig. 4. Electronic hub circuit comprising two lines and a loop leg

at the board. Some time later, when the network is idle, the temporary arrangements are replaced by more permanent cross-connections between facilities or legs. Patches are also employed to introduce spare facilities quickly for the purpose of replacing assigned facilities on which trouble has developed. No effort is made to clear these troubles until the assigned facility that is in trouble has been released from service. The face of the board is ordinarily free from patching cords except when required for temporary rearrangements or for testing.

The ease with which these temporary patches may be used to expedite building up new networks and to modify existing layouts is one of the important features of the serviceboard. It is made feasible by the use of the electronic hub circuit because network changes do not affect the hub circuit currents and because all facilities (line or loop) terminate in a uniform manner. Another important advantage of the hub circuit is that transmission through it is practically independent of the number of legs connected to the hub within wide limits. A hub can accommodate from 12 to 20 legs, the exact number depending on the ratio of electronic coupling units to loop repeaters used, the capacity of the hub leads to ground, and whether or not regeneration is provided.

For hub circuit operation, each line or loop facility terminates in a legs relay circuit which is associated with a legs jack appearance in the face of the service position (see Fig. 6). An electronic regenerative repeater is terminated in a similar manner except that it is associated with a hub relay and legs jack circuit. The legs jack circuits of all line and regenerative repeater facilities have jack appearances in all service positions; only spare loops have these multiple jack appearances. All multiple legs jack appearances are equipped with idle indicator lamps to indicate whether the facility is busy or idle. The legs jacks permit the associated facility, when spare, to be patched for service or for testing and maintenance.

When a facility is assigned for service, its send and receive legs are cross-connected to the *SH* and *RH* leads of a hub circuit and its relay circuit is cross-connected to a jack associated with the hub circuit. Line or loop legs are associated with T jacks and the regenerative repeater with the hub jack. A hit indicator lamp is located above each T jack of the hub circuit to assist in the location of troubles. A flashing lamp indicates that the leg is sending to the hub and all other lamps of the hub should remain out. Momentary or steady lighting of lamps associated with legs which are receiving from the hub is an indication of trouble.

Leg-patching cords are employed for most of the patching done in conjunction with hub circuits. The cords are located in the face of the board where they are readily accessible, and 40 cords are provided in each service position. Service patches perform three principal functions; namely, the addition of spare facilities, the substitution of spare for assigned facilities, and the building up of new networks on a temporary basis. An addition of a line or loop leg is made by patching the legs jack of the spare facility to the hub jack.

A spare line, loop, or regenerative re-

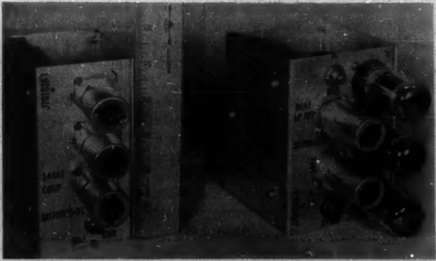


Fig. 5. Electronic coupling unit and electronic loop repeater

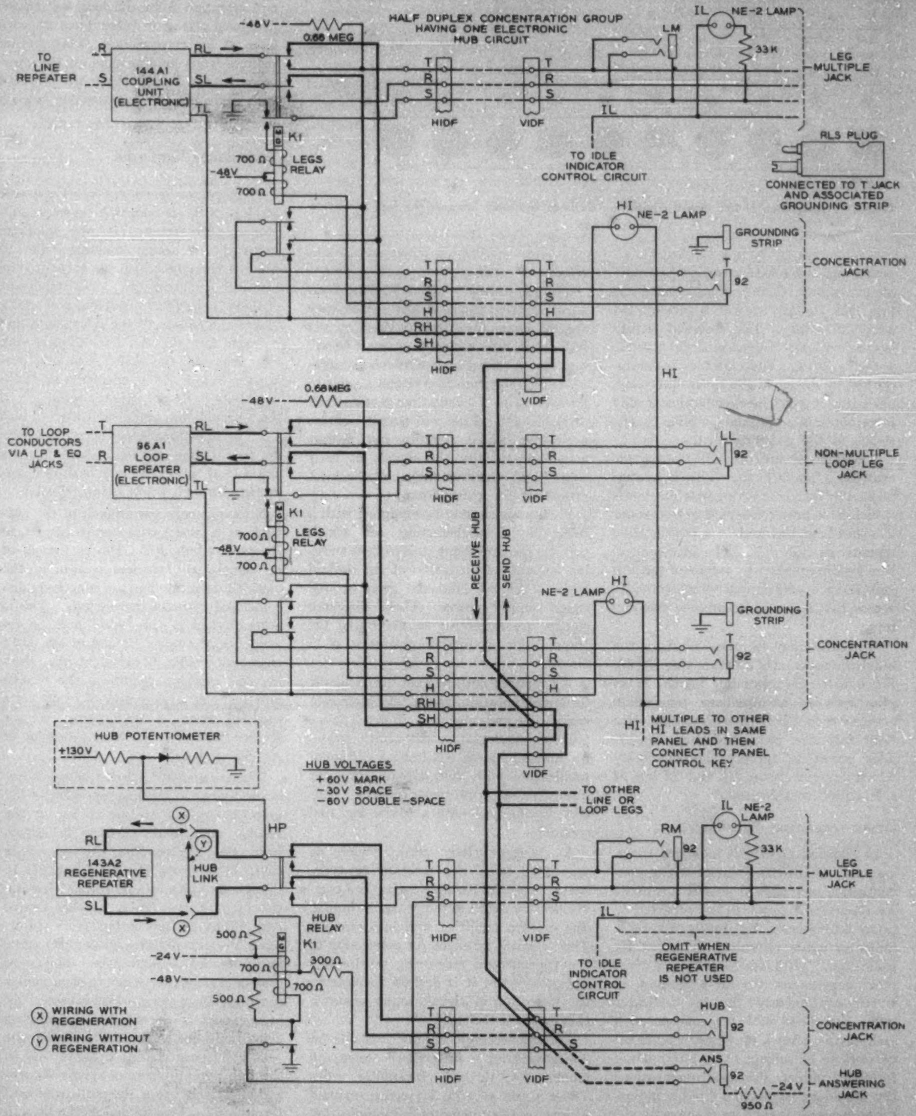

HALF DUPLEX CONCENTRATION GROUP HAVING ONE ELECTRONIC HUB CIRCUIT
-48V
0.68 MEG
TO LINE REPEATER
R
S
144A1 COUPLING UNIT (ELECTRONIC)
RL
SL
TL
K1
LEGS RELAY
700 Ω
-48V
700 Ω
T
R
S
H
RH
SH
HIDF
VIDF
LM
IL
NE-2 LAMP
33K
LEG MULTIPLE JACK
IL
TO IDLE INDICATOR CONTROL CIRCUIT
RLS PLUG
CONNECTED TO T JACK AND ASSOCIATED GROUNDING STRIP
HI
NE-2 LAMP
GROUNDING STRIP
T
92
CONCENTRATION JACK
HI
-48V
0.68MEG
TO LOOP CONDUCTORS VIA LP & EQ JACKS
96A1 LOOP REPEATER (ELECTRONIC)
LL
92
NON-MULTIPLE LOOP LEG JACK
RECEIVE HUB
SEND HUB
NE-2 LAMP
HI
GROUNDING STRIP
CONCENTRATION JACK
HI
MULTIPLE TO OTHER HI LEADS IN SAME PANEL AND THEN CONNECT TO PANEL CONTROL KEY
HUB POTENTIOMETER
+130V
HUB VOLTAGES
+60V MARK
-30V SPACE
-80V DOUBLE-SPACE
TO OTHER LINE OR LOOP LEGS
HP
143A2 REGENERATIVE REPEATER
X
Y
HUB LINK
RM
92
IL
NE-2 LAMP
33K
LEG MULTIPLE JACK
OMIT WHEN REGENERATIVE REPEATER IS NOT USED
TO IDLE INDICATOR CONTROL CIRCUIT
HUB RELAY
K1
500 Ω
-24V
700 Ω
300 Ω
-48V
700 Ω
500 Ω
(X) WIRING WITH REGENERATION
(Y) WIRING WITHOUT REGENERATION
HUB
92
CONCENTRATION JACK
ANS
92
-24V
950 Ω
HUB ANSWERING JACK

Fig. 6. Method for interconnecting the legs and the regenerative repeater or hub link for assigned hub circuit service. Associated jacks and relays provide for adding legs or for making substitutions for assigned legs or regenerative repeaters by means of patches

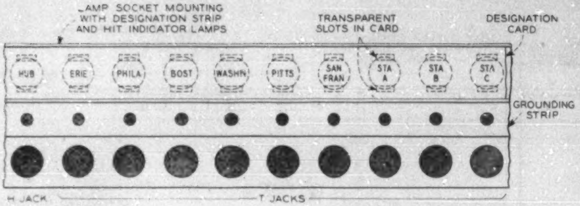


Fig. 7. Appearance of hub circuit comprising six lines and three loops in the face of service position

peater may be substituted for the equivalent assigned facility by patching its legs jack to the T or hub jack of the assigned facility. The assigned facility is automatically released from the hub with the patch. The switching is accomplished by means of the leg or hub relay, operation of which is controlled marginally from potential or ground supplied by the sleeve of the patching cord.

A hub circuit may be built up temporarily on a patch basis. This is accomplished by interconnecting patching cords which have been connected to the proper line and loop legs and to a regenerative repeater or hub link. The interconnection is accomplished by means of the hub jack in each cord circuit which provides access to the two transmission conductors.

A facility may be released from the hub by connection of a release plug (*RLS*), Fig. 6, to the corresponding T jack. This plug includes an auxiliary pin which engages with a hole in a grounding strip built into the jack mounting, Fig. 7. This serves to operate the legs relay which disconnects the *RL* and *SL* legs of the facility from the hub.

Other Improved Features

In addition to the hit indicator lamps associated with the T jacks of the concentration jack circuit, provision is made for observing hits which are reported to occur infrequently, by means of trouble indicator cords providing locked-in hit indication. With trouble indicator cords it is unnecessary for an attendant to watch continuously the hit indicator lamps associated with the concentration jack circuit when a customer reports his network to be subject to an intermittent form of trouble. For such observations the attendant connects a trouble indicator cord circuit to each leg of the network; and each position is equipped with six trouble indicator cord circuits, for this purpose. This permits the attendant to determine on which leg of the network trouble is occurring.

The serviceboard includes jack appearances of teletypewriter test signals with special means for patching them to any hub or facility where they may be required. A test signal cord circuit provides for the connection of test signals to lines or loops for long-time testing to a distant point. This cord circuit includes a feature which permits the distant attendant to stop the signals at their source and simultaneously call the local attendant for consulation.

Each position may be equipped with a telegraph signal-distorting set which permits the attendant to introduce varying amounts of distortion of any desired type on signals obtained from a test signal supply source. These distorted signals are employed to determine the transmission distortion tolerance of a circuit.

A telephone cord and set are available for communication over a telephone subscriber line and over intraoffice telephone trunks. A telegraph cord circuit provides for manual Morse monitoring and communication with hub circuits or over interoffice telegraph trunk circuits. This cord is equipped with a Morse key and sounder.

A teletypewriter cord, which is associated with the position teletypewriter, provides for monitoring and communication over interoffice trunk circuits, line or loop facilities, and hub circuits. This circuit is arranged for connection to the transmission measuring set by key operation when it is desired to measure the distortion of signals under observation by the cord circuit.

Two connecting cords are available for interconnecting interoffice telegraph trunks for through interoffice calls. These cords provide for either manual Morse or teletypewriter communication under key control. Each cord includes a disconnect feature responsive to ringing signals, which permits either the calling or called attendant to disconnect the cord from both trunk circuits electrically without attention from the local attendant. The local attendant takes down the cord later at his convenience. This feature reduces trunk-holding time on the interoffice trunks, all of which are provided with multiple jack appearances in each serviceboard.

Techniques Employed

Prior to the development of the new serviceboard, all telegraph repeaters employed polar relays. It was therefore necessary to design coupling units to connect these relay-type repeaters to the new hub arrangement. Most important of these is the *144A1* coupling unit which makes connection to the relays in a line repeater (see Fig. 5). An all-electronic loop repeater, the *96A1*, was also designed, Fig. 5. The operation of these two units is described with reference to the 3-branch hub circuit indicated in Fig. 4. Two line branches are shown connecting to the hub through coupling units and a local subscriber loop is shown connecting through a *96A1* loop repeater.

Normally only one branch of the network at a time passes signals inward to the receive hub, *RH*. During periods of no signals, all branches remain in the marking condition and are effectively disconnected from the receive hub. This is accomplished in the *144A1* unit by reverse-biased diode *CR1* and in the *96A1* repeater by tube *V2* being cut off. During this marking condition, the receive hub rests at a potential of +60 volts supplied from a terminating voltage divider (hub potentiometer) consisting of *R1*, *R2*, and *CR2*.

When an inward spacing signal is imposed by one of the line repeaters, −130 volts is applied to the *R* leg of the *144A1* unit and current is drawn from the receive hub through diode *CR1*. This drops the hub potential to the spacing value of −30 volts. Opening the loop of the *96A1* repeater causes tube *V2* to conduct current from the receive hub and likewise lower the receive hub potential to −30 volts. Thus, any branch of the network can cause a similar spacing voltage swing which is essentially independent of the number of branches being multipled at the hub. During the −30-volt space condition diode *CR2* in the hub potentiometer is reverse-biased, and effectively opens the path to ground through *R2*. This minimizes the drop in hub resistance during the space condition and enables the −30-volt potential to be reached with less current flow in the transmitting leg. The more constant hub resistance also results in a nearly symmetrical hub signal wave and minimizes signal bias when the wave becomes rounded by capacitance to ground in a large hub circuit or concentration group.

If no regeneration is required, the receive and send hubs are directly interconnected by means of a hub link. When the two hubs are interconnected through a regenerative repeater, the same marking and spacing potentials are produced at the send hub. However, as a result of the regeneration, the signal train is delayed by one-half signal element.

The outward paths are coupled to the send hub through high resistance and are essentially voltage-operated. Consequently the voltage swing of the hub does not change materially as the number of outward paths is increased. In the *144A1* coupling unit, tube *V1* in the outward path serves to drive the send relay *S* of the line repeater in accordance with the signal on the send hub. In the *96A1* repeater, tube *V3* in the outward path repeats the signal by opening and closing the loop circuit.

ol. 74
. 1
955

The resistance of the hub circuit is about 2,000 ohms, low enough to avoid excessive crosstalk into other hubs. This permits the use of up to 10,000 feet of switchboard wire in setting up a hub connection without excessive rounding of the signal wave. The number of branches multipled at a single hub is, however, limited to about 20 because the connections are not infinite in resistance. Where a larger concentration group is required, more than one hub may be set up with interhub repeaters interconnecting them.

The *144A1* coupling unit includes a directional control feature which prevents the flow of signals outward from the send hub to the line repeater which originated them. This control includes a bistable circuit triggered by the combination of voltages at points *a* and *b*. When a space is sent through the *144A1* to the receive hub, both points swing negatively. If, however, a space is sent to the hub from some other branch, only point *b* swings negatively. This difference causes the bistable circuit to be set to indicate the direction of the last pacing pulse on the receive hub. When an inward flow is indicated, point *c* is clamped in a marking condition and outward flow of signals is blocked.

If two subscribers send simultaneously, or if one purposely breaks the sending of another, two space signals may be impressed on the receive hub. This is known as a double space and results in a hub potential of −60 volts. The directional control is arranged so that this condition will unblock the outward paths of those branches sending the spaces to the hub so that all branches receive a space condition. This has been called a double-space by-pass and insures that all spaces sent inward from one branch are transmitted outward on the other branches and are never lost because of a prior space condition on the hub. This makes simultaneous sending immediately evidenced by garbled local copy, and gives a quicker "break."

As mentioned, the *96A1* loop repeater transmits to the loop on an open-and-close basis. This is made practical by the high dynamic plate resistance of tube *V3*, which sends a symmetrical current wave into the loop even when it consists of 25 miles of cable pair. This results in unbiased signals at the subscriber's equipment without need of any adjustments.

Transmission from the subscriber is also on an open-and-close basis but the signal is received at the *96A1* on a voltage basis. In the normal marking condition of the loop, the potential at the plate of *V3* is −50 volts. When the loop is opened at at the subscriber's unit, the plate falls to nearly −130 volts. This voltage swing is inverted and coupled to tube *V2*, which in turn transmits to the receive hub. When a space is sent toward the subscriber, the plate potential of *V3* swings postitive and thus a mark condition is held toward the hub. The high dynamic plate resistance of *V3* causes the voltage wave received from the subscriber to be symmetrical, and no signal bias occurs as the length of loop is increased. Furtherfore, no inductive loading at the subscriber's unit is necessary as when receiving on a current basis.

The *96A1* repeater also has a directional control feature similar to that in the coupling unit. This serves to hold *V3* conducting when sending toward the hub. This directional feature does not need to recognize a double space since

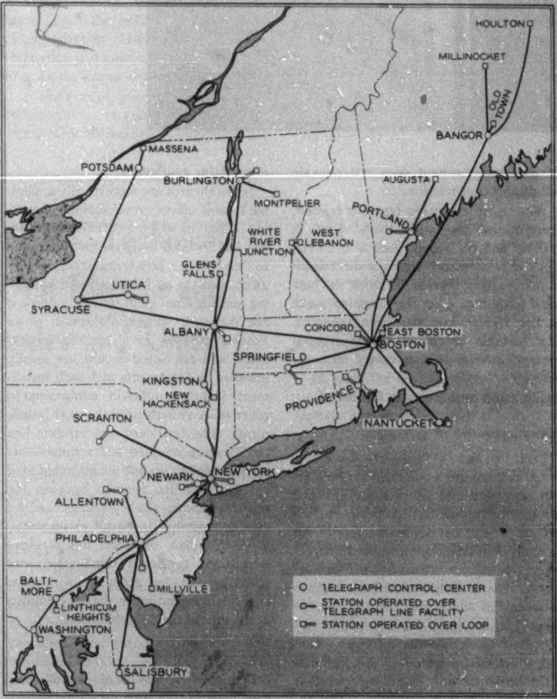


Fig. 8. A complex half-duplex private-line network

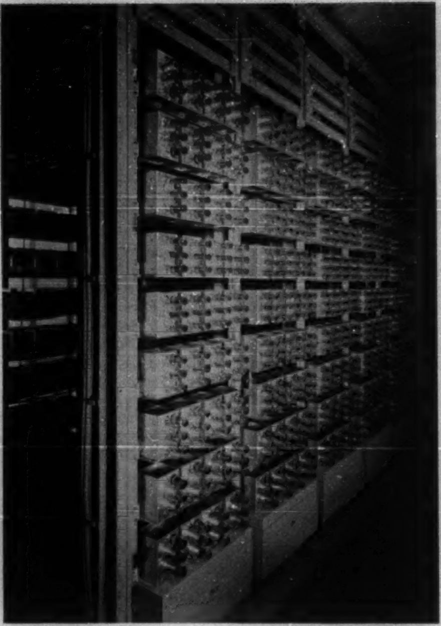


Fig. 9. Bay-mounting arrangements for electronic loop repeaters

the subscriber's loop does not provide full-duplex transmission.

When a regenerative repeater is used, there are intervals when the send hub is spacing and the receive hub is marking. Thus it is necessary that the directional control feature be able to block the outward path even after the incoming space to the receive hub has been removed. The bistable circuit in the directional control provides this necessary memory.

Application and Results Achieved

APPLICATION

Each line or loop facility requires electronic terminating equipment for hub circuit operation and hence a serviceboard installation involving 100-per-cent hub circuits would cost somewhat more than an equivalent telegraph testboard. In replacing existing testboards with serviceboards it is desirable to strike a balance between this added cost and the benefits derived from operating economies and improved service to the customer. One step which has been made in this direction is the introduction of direct-leg operation for simple, 2-branch networks. This results in material savings in electronic terminating equipment.

Serviceboards are more readily applicable to new offices, or relocations of existing offices, in which case new line facilities are installed designed for connection to electronic hub-terminating equipment, and the added expense of modifying existing equipment for this purpose is not incurred. When conversion of an existing office is involved, the serviceboard and associated circuits are usually installed in increments and each installment includes sufficient equipment to cover office growth until the next installment. The testboard and existing line facilities are retained and arrangements are provided to permit the interchange of old and new line facilities between the testboard and serviceboard so as to afford greater flexibility in circuit assignments. With this plan, the older line facilities with their testboard terminations are scheduled to be replaced with new facilities terminating in the serviceboard at the time of their obsolescence.

RESULTS ACHIEVED

Complete serviceboard installations have been in operation for about 2 years in new offices at Oakland, Calif., Portland, Ore., Chicago (Stewart), and New York (West 50th Street). Operating experiences in these offices have been most gratifying. Troubles can be directionalized, sectionalized, and removed more quickly than with testboards.

An improved feature of importance which has not been mentioned previously is the hub circuit jack field arrangement. This is so simple that an attendant can determine at a glance the layout of a complicated network. The simplicity of a concentration jack circuit appearance in the service position is evident from Fig. 7. With testboard operation the attendant must read the designation cards over many jacks and in some cases refer to layout sketches to obtain this information. Another factor contributing to the difficulty in determining the nature of a circuit layout in a testboard office is the relative nonuniform manner in which networks are interconnected in these offices. The same group of facilities may be interconnected in a number of ways in a testboard, whereas there is little, if any, variation permissible in a hub circuit network.

These improvements provide better service and permit more circuits to be handled per attendant. However, the maximum operating efficiency obtainable with serviceboard arrangements cannot be realized until serviceboards have been installed in all of the larger offices. This is because interoffice teamwork is involved in performing certain operations and if one office is appreciably slower than another, the higher efficiency of the faster office is not fully effective.

The training problem has been simplified because the uniform termination of facilities in serviceboard offices has permitted the adoption of standard procedures for testing and patching. In the case of testboard operation, layouts are so varied that it is difficult to establish standard procedures for testing and patching operations to cover all network arrangements, and each attendant must develop his own technique from knowledge based on previous experience with similar networks.

The West 50th Street office in New York is considered a typical serviceboard office. It includes five service positions which are shown in Fig. 3 and two facility positions which are currently used to interconnect 600 line facilities, 520 2-way loop facilities, and 113 1-way loop facilities to form 241 TWX trunk terminals, 23 TWX through circuits, 184 private-line terminal circuits, and 42 private-line networks arranged for hub circuit operation. Each TWX or private-line terminal circuit requires the interconnection of a line facility and either a trunk terminal or a loop. The number of line and loop facilities required for the private-line hub circuit networks will vary from 3 to usually not more than 20 facilities depending upon the circuit complexity. The average for the 42 circuits at West 50th Street is 2.1 line facilities and 2.4 loop facilities. A typical complex half-duplex network terminating in a hub circuit at West 50th Street, New York, is shown in Fig. 8. This layout includes six legs at New York and nine at Boston.

HEAT PROBLEM

A new problem which has been encountered in connection with the development of the telegraph service board is that of removing heat dissipated by the small, compactly mounted electronic equipment. The manner in which these demountable units are bay-mounted is shown in of Fig. 9, which shows several bays of loop repeaters. Owing to the relatively small amount of heat dissipated by the older relay-type telegraph repeaters and the relatively large areas which they occupy in the bay, heating heretofore has been no problem in telegraph offices. The heating problem has two important aspects, namely, the protection of the electronic apparatus components from excessive temperature rise and the prevention of excessive operating room temperatures.

Several steps have been taken to minimize excessive heating. It has been determined by field experience that the maximum allowable heat dissipation per bay should be limited to 1,000 watts. When several adjacent bays are occupied by hot equipment, it is desirable either to provide greater separation between these bays or to maintain the total dissipation per bay well below 1,000 watts. This is necessary in order to protect apparatus units such as resistors, varistors, and condensers.

Certain design features have been embodied in the electronic units in order to minimize the heating problem. The units, which are all of a demountable type, are designed so that the vacuum tubes are located on the outside where heat dissipation is assisted by room air convection currents and where dissipation by the tubes is less likely to affect other apparatus located inside the unit. The unit enclosures are left open at the top and bottom and the units are arranged for closely adjacent vertical mounting so as to provide a chimney effect for the creation of a cooling air draft.

In some cases, where there are heavy concentrations of this electronic equipment, air-conditioning equipment has been installed to supplement the ventilation system normally available in the building. The air-conditioning units are commonly located along one wall of the room and the cool air is distributed throughout the room by means of ducts. The present need for such air-conditioning indicates the need for further development to reduce the heat generated in electronic units of this type.

Conclusions

The new serviceboard arrangements are an important step forward in the development of means for providing good service and for meeting the present complex requirements for teletypewriter private-line and TWX services. They also include facilities suitable for meeting future needs, considering the present trend toward higher transmitting speeds and increased complexity of teletypewriter and similar services such as teletypesetter, transmission of punched-card data, etc.

References

1. A NEW ELECTRONIC TELEGRAPH REGENERATIVE REPEATER, B. Ostendorf, Jr. *Electrical Engineering*, vol. 69, March 1950, pp. 237–40.

2. A CARRIER TELEGRAPH SYSTEM FOR SHORT-HAUL APPLICATIONS, J. L. Hysko, W. T. Rea, L. C. Roberts. *Electrical Engineering*, vol. 71, July 1952, pp. 625–30.

Recommendations of IEC Technical Committee No. 24 on Electric and Magnetic Magnitudes and Units

J. J. SMITH
MEMBER AIEE

A MEETING of Technical Committee (TC) 24 on Electric and Magnetic Magnitudes and Units (EMMU) was held at Philadelphia, Pa., on September 10 and 11, 1954, during the Golden Jubilee of the International Electrotechnical Commission (IEC). It should be understood that the recommendations made by this committee, and given in the following, have been referred to the national committees of the various countries for an official vote under the 6 months' rule. Thus they are not as yet official decisions. However, since each recommendation was approved by a large majority of the delegates present from the various countries, it seems reasonable to expect that they will be approved by the national committees.

It is hoped that the reader will bear in mind that each action in the following account is the best that can be agreed upon at present and should be considered as the most practical clarifying step that can be taken now. Further progress depends on: 1. a more nearly perfect understanding of the phenomenon involved, and 2. a better agreement on the meaning of the technical words or language used in the exchange of different viewpoints. Both of these will come with time. The recommendations are listed in the following.

Sign of Reactive Power

In 1935, the IEC adopted the convention that the reactive power flow to an inductive receiver be negative. In 1946, an AIEE committee undertook a survey of the prevailing practice in the United States. They found that it varied, but that most power companies assumed the sign of reactive power to an inductive receiver to be positive. In accordance with this, the AIEE changed the standard to accord with this practice. This was brought to the attention of the IEC at

---

Paper 55-49, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing December 8, 1954.

J. J. SMITH is with the General Electric Company, Schenectady, N. Y.

the Paris meeting in 1950 and it was suggested that other countries review their usage. Since then, many of the countries have found that they are also using the sign of reactive power to an inductive receiver as positive. It is hoped that the question will be clarified by the following resolution, which was adopted, and with which the latest American Standard is in accord. Resolved that:

1. Reactive power flow to an inductive receiver be positive.
2. Only one kind of reactive power be considered and defined. This means that no qualifying adjective should be used with the term "reactive power."
3. Reactive power be completely specified by stating its amount and direction of flow.

Name for MKS or Giorgi System

It was recommended that the system which uses the fundamental units meter, kilogram, and second (MKS) be officially known in the IEC as the Giorgi system.

It was recognized that the system which uses the fundamental units, meter, kilogram, second, and ampere is referred to by many as the MKSA system. This term is similar to the term centimeter-gram-second (CGS) system.

Assignment of Names "Magnetic Polarization," "Intrinsic Induction," and "Magnetization"

The magnetic induction B at a given point in a magnetic field is represented by a vector the magnitude of which is equal to the maximum limit of the time integral of the electromotive force generated in a small plane-conducting loop brought in from a region of zero magnetic field divided by the area, as the area approaches zero in a macroscopic sense. The maximum limit occurs for a particular orientation of the loop at the given point. The direction of the vector is normal to this loop.[1]

The magnetic field strength H at a given point is represented by the current per turn multiplied by the number of turns in a small elongated uniformly wound solenoid divided by its length when the solenoid is centered at the given point and is oriented and adjusted so as to counterbalance magnetic effects at that point.[2]

In vacuum the ratio $B/H=\mu_0$ is called the permeability of free space. In any other medium the ratio $B/H=\mu$ is called the permeability of the medium.

It was recommended that in a medium other than vacuum the term $B-\mu_0H$ be called "magnetic polarization" or "intrinsic induction," and that $(B/\mu_0)-H$ be called "magnetization." These terms are used by many authors and a name is needed for them. There was considerable discussion as to which of these two terms should be given preference in the recommendations. The views of the national committees will be obtained on this. Note: In the United States the symbol μ_v is used instead of μ_0.

Name for Unit of Magnetic Induction in Giorgi System

It was recommended that the name "tesla" be given given to the unit of magnetic induction (the weber per square meter) in the Giorgi system.

Similarly, in the centimeter-gram-second electromagnetic (CGSM) system, there is the gauss, the unit of magnetic induction, which is equal to the maxwell per square centimeter.

Rationalization of the Electromagnetic Field Equations

At the IEC 1950 Paris meeting TC 24 recommended that the MKS system be used in the rationalized form. At that time, questions were raised as to whether rationalization by change of quantities, or change of units was meant. Since then much discussion has been devoted to the subject. The following was recommended at the 1954 meeting:

"The rationalization of the equations of the Electromagnetic Field recommended by the I.E.C. is characterized by the following principal equations whose form remains the same whether the symbols are regarded as representing physical quantities coming into play or are regarded as representing their numerical values.

$$\oint H\,ds=\Sigma I$$
$$\phi=\int\int B\,dA$$
$$B=\mu H$$
$$\int\int D\,dA=Q$$
$$\psi=\int\int D\,dA$$
$$D=\epsilon E$$

"These equations apply to isotropic media. The total current (ΣI) includes in the general case the displacement current.

"The permeability of free space μ_0 has the numerical value $4\pi\times10^{-7}$, where the unit is the henry per meter. The permittivity of free space ϵ_0 has the value of $(1/c^2\mu_0)$ or approximately the numerical value 8.85×10^{-12}, where the unit is the farad per meter." (The table of references for the quantities used in the foregoing equations is shown below.)

Relation of Ampere Turn per Meter to Oersted

It was found that in the usage of various writers the relation between the ampere turn per meter and the oersted was apparently not always the same. To secure uniformity, the following resolution was adopted:

> A current of ten amperes in a wire of one centimeter length of arc, of a circle of one centimeter radius, contributes at the center of the circle a magnetic field strength of one oersted.
>
> A uniform elongated solenoid having one turn per meter axial length in which there is a current of one ampere produces at the center of the solenoid a magnetic field strength of one ampere-turn per meter.
>
> The contribution to the magnetic field strength in the first case is equal to $1{,}000/4\pi$ times the magnetic field strength in the second case.

Here the first paragraph of the resolution uses the arrangement by which the abampere was defined by means of the oersted in the British Association Reports on Electrical Standards.[3] The second paragraph uses the arrangement by which the ampere-turn per meter is defined in the IEC vocabulary.[2] The third paragraph gives the relation between the two magnetic field strengths.

It will be noted that actual numerical examples are used in the resolution. This was done to avoid the use of symbols, since these are also used in different senses by different authors. It is hoped that the resolution will produce uniformity in the relation between the ampere-turn per meter and the oersted as used in future.

Table of Reference for the Quantities in the Above Equations

Name	Symbol	Unit of the Giorgi System	IEC Classification Number
Element of Length	ds	Meter	05-35-050
Electric Charge	Q	Coulomb	05-15-010
Electric Field Strength	E	Volt per meter	05-15-045
Permittivity	ϵ	Farad per meter	05-15-120
Electric Displacement Density	D	Coulomb per sq meter	05-15-130
Displacement Flux	ψ	Coulomb	05-15-135
Electric Current	I	Ampere	05-20-040
Magnetic Field Strength	H	Ampere turn per meter	05-25-022
Magnetic Induction	B	Weber per sq meter	05-25-030
Magnetic Flux	ϕ	Weber	05-25-035
Permeability	μ	Henry per meter	05-25-045

Revision of Definitions in Second Edition of International Electrotechnical Vocabulary Within Scope of TC 24

Because of certain points which were discussed in the previous resolutions in connection with definitions prepared by TC 1, TC 24 agreed that it should consider carefully the definitions which were closely allied with its activities. Some definitions to which attention had already been called were those of reactive power, complex power, permeability, and permittivity.

Comparison of Systems with Three and Four Dimensions

At the 1950 Paris meeting the IEC adopted the ampere as the fourth preferred fundamental unit in the Giorgi system. The International Union of Pure and Applied Physics (IUPAP) adopted the ampere as the fourth fundamental unit in the Giorgi system in 1951.

Thus the MKSA system is based on four fundamental units (meter, kilogram, second, and ampere). The two CGS systems (electrostatic and electromagnetic) are based on three fundamental units (centimeter, gram, and second). In the CGS electromagnetic system, the permeability of free space is assumed to have no dimensions and its numerical value is considered equal to 1. In the CGS electrostatic system, the permittivity of free space is assumed to have no dimensions and its numerical value is taken to be equal to 1. It was agreed that the relations between these three systems, each of which is based on a different assumption, should be studied and clarified. A memorandum by Dr. H. Konig and Dr. U. Stille has been prepared on this subject and will be circulated to the national committees for study.

Preparation of Document Containing all Decisions of TC 24 Since Founding of IEC

It was agreed that such a document should be prepared by the secretariat so that all the decisions of TC 24 from its beginning to date would be readily available.

It was pointed out that reference to other sources will probably be necessary later,[1,4,5] and that as a first step collecting the IEC decisions was desirable. The work on electric and magnetic magnitudes and units is carried on in this country by the American Standards Association Committee C61.

References

1. British Association Reports on Electrical Standards. Cambridge University Press, Cambridge, England, 1913. See p. 520 for a similar way of defining magnetic induction.

2. International Electrotechnical Vocabulary. International Electrotechnical Commission, Geneva, Switzerland, 1954, IEC classification no. 05-25-022.

3. See reference 1, p. 102.

4. Reports on Symbols, Units and Nomenclature of the International Union of Pure and Applied Physics. *Physical Society*, London, England, 1935.

5. Systems of Electric and Magnetic Units. *Bulletin No. 93*, National Research Council, National Academy of Sciences, Washington, D. C., 1938.

Eddy-Current Mutual-Inductance Transducers With High-Conductivity Reference Plates

HORACE M. JOSEPH
MEMBER AIEE

NATHAN NEWMAN
NONMEMBER AIEE

Synopsis: The design considerations of mutual-inductance devices for use in noncontacting measurements of distance are studied with special emphasis on the eddy-current type utilizing "perfectly" conducting metal reference plates of infinite extent. The force on the plate due to the exciting current, typical schematic circuits used to detect the results, and the effects of rapid motion of the gauge are given. Since actual reference plates are not perfectly conducting nor infinitely extensive, some indications of the effects of finite conductivity and limited extent of reference plates are included.

THIS paper analyzes the general performance characteristics of mutual-inductance transducers. This analysis is an outgrowth of the use of the radio-frequency mutual-inductance transducer invented by M. L. Greenough. Several development projects have utilized the micrometer for specific purposes, but each design analysis was restricted to the specific application. Accordingly it was believed desirable to establish general design criteria for these mutual-inductance transducers.

Noncontacting Distance Gauges

The principal methods in use for measuring small distances by noncontacting means may be classified as optical, acoustical, and electrical. Optical methods are in wide use, but an optical path is not always available for measuring purposes. In addition, if the indication by an optical method is to be used for control purposes, the amount of energy available is so low as to require considerable amplification. Acoustical methods which are typified by sonar, also suffer from the inability to provide high-energy signals directly.

Gauges using electrical phenomena may be classified into propagation-echo types exemplified by radar (which are not suitable for use in measuring small distances) and other electrical methods that use a variation of one or more of the basic electrical parameters (resistance, capacitance, or inductance) as a function of the distance or displacement to be measured. Change in electrical resistance with displacement is not readily adaptable to noncontacting methods. Variation in capacitance is limited to the measurement of very small distances. Also it is necessary that the dielectric constant of the material between the objects whose capacitance is to be measured remain fixed. The

---

Paper 55-4, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted December 11, 1953; made available for printing October 27, 1954.

Horace M. Joseph is with ACF Electronics, Alexandria, Va., and Nathan Newman is with the National Bureau of Standards, Washington, D. C.

This paper reports the results of an investigation conducted by the Electronic Instrumentation Section of the National Bureau of Standards under a program of Basic Instrumentation Research and Development sponsored by the Office of Naval Research, the Air Research and Development Command, and the Atomic Energy Commission.

The valuable suggestions of Dr. L. S. G. Kovasznay of The Johns Hopkins University are gratefully acknowledged.

Table I. Terminology

Method of Varying Self- or Mutual Inductance	Designation of Method
Variation of effective permeability by relative motion of ferromagnetic metal in vicinity of coil(s)	Variable-mu inductance
Relative motion nonmagnetic metal plate in vicinity of coil(s) (typified by Fig. 1)	Eddy-current inductance
(a) Perfect-conductor case	(a) Perfect conductor
(b) Resistive-conductor case	(b) Resistive conductor
Relative motion of coils (or two parts of one coil) (typified by Fig. 2)	Movable-coil inductance

Table II. Comparison of Movable-Coil- and Eddy-Current-Inductance Distance-Measuring Devices

Consideration	Movable-Coil Inductance; See Fig. 2	Eddy-Current Inductance; See Fig. 1
Sensitivity at a given separation	Less than eddy current	More than movable coil
Coil support	Separate support for each coil	Common coil support
Leads	Flexible leads for movable coil Fixed leads to stationary coil	Fixed leads to both coils
Changes in mechanical alignment:		
(a) Lateral	Sensitive to lateral shift	Insensitive to lateral shift
(b) Tilt	Less sensitive than eddy current	More sensitive than movable coil
Fixed component of response	None	Must be balanced out

output of such instruments occurs at a low-energy level, and, except for very small separations, the variation is a small percentage of a fixed capacitance consisting largely of that of the leads. In a capacitance distance-measuring gauge, it is necessary either to make a connection to the reference plate or to see that the reference place is well insulated. Nevertheless this type of distance gauge has been used extensively.

Inductance-Type Distance Gauges

The remaining electrical parameter that may be varied by displacement is inductance, either self or mutual. This variation may be accomplished in a variety of ways. Table I gives the terminology used. In Table I the eddy-current inductance method is subdivided into two cases: the perfect conductor and the resistive conductor. The distinction depends upon the depth of eddy-current penetration below the surface of the plate. If this depth is negligible in comparison with all dimensions of the system, the plate will be considered as a perfect conductor. What is considered a negligible depth of penetration will be discussed later.

The following observations about the resistive conductor may be made. The effect of a resistive plate in changing a mutual inductance is less than that of a perfectly conducting plate. The greater the resistivity of the plate material, the less is the effect on mutual inductance. Furthermore, if the current at the back surface of the plate is not negligible (a thin resistive plate), the effect on mutual inductance is less than it would be with a thick resistive plate.

Self-inductance and mutual inductance are manifestations of the same phenomenon of flux linkages. For purposes of calculation, self-inductance can be considered as the mutual inductance between two identical coils in the same position. In a practical circuit, the voltage across a self-inductance must include a resistive voltage drop due to the dissipation of the winding. Since this is not necessary when using an open-circuit secondary coil coupled by mutual inductance, this type of circuit is generally more practical.

The three ways of varying inductance, listed in Table I, are quite similar and have many advantages and disadvantages in common. Some disadvantages are:

1. The circuits may be subject to resonant effects that, without careful design, may lead to instability.
2. The range of distances which may be measured is limited to those comparable to the coil diameter(s).
3. Measurements may be adversely affected by the presence of extraneous metal in the vicinity.

The advantages common to all of these inductance-type distance gauges include:

1. High-energy level output.
2. Small mechanical forces on the reference plate.
3. Stability of response.
4. Applicability to a wide range of displacements (down to as small as 1 microinch).
5. Independence of dielectric constants of materials in the gap between the fixed coil and the reference object.

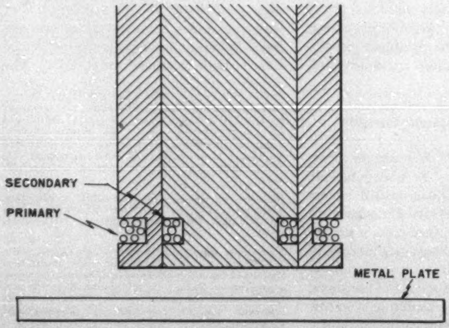


Fig. 1. Mutual-inductance probe with both coils relatively fixed on same form

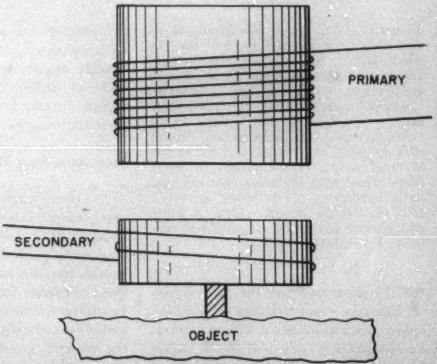


Fig. 2. Mutual-inductance probe with secondary attached to object and movable

6. Output little affected by normal lead capacitances or inductances.

The variable-mu-inductance devices, for measuring distances, are different from the eddy-current and movable-coil types in that large variations in effective permeability are easily obtained; hence the principle involved can be used to make instruments that are very sensitive to small static displacements. These devices can be used at ordinary power frequencies. Actually, they require use of a relatively low frequency which makes them unsuitable for measuring very rapid displacements. The higher magnetic fields necessary for a given voltage output cause comparatively large forces to be exerted on the reference plate. In addition variable hysteresis tends to cause some errors.

The eddy-current and movable-coil distance-measuring devices are very closely related. The same type of calculations may be used to determine the changes of mutual inductance. In a movable-coil device the output is the direct (inverse) function of separation distance. In the eddy-current device the output is the difference between a fixed quantity and an inverse function of separation distance. Accordingly, for large distances compared to the radius of the coil, the movable coil is preferable. These features are compared in Table II.

This paper relates primarily to eddy-current-inductance distance-measuring gauges. Inasmuch as the process of obtaining the solution to the eddy-current method involves the calculation of the mutual inductance between pairs of coils, a considerable part of the information included is applicable to the movable-coil system. The variable-mu method will not be treated, nor will the effects of heating of the reference plate be considered.

Nomenclature

$j=\sqrt{-1}$
$\omega=2\pi f$ (f = frequency, cycles per second)
$I_{1\,\max}$ = maximum value of primary current, amperes
I_1 = rms value of primary current, amperes
$I_{1\,\max}'$ = maximum value of current in image of primary coil, amperes
I_1' = rms value of current in image of primary coil, amperes
$R_{2\max}$ = maximum value of open-circuit voltage across the secondary coil, volts
E_2 = rms value of open-circuit voltage across the secondary coil, volts
M_{12} = mutual inductance between primary and secondary coils, henrys
M_{12}' = mutual inductance between image of the primary coil and the secondary coil, henrys
n_1 = number of turns on the primary coil
n_2 = number of turns on the secondary coil
A = average radius of the primary coil, centimeters (as in Fig. 6)
a = average radius of the secondary coil, centimeters (as in Fig. 6)
x = distance between the planes at the centers of the two coils under consideration, centimeters
y = a function of a/A and x/A
d = distance between a plane at the center of the secondary coil and the reference plate, centimeters (as in Fig. 6)
b = distance between planes at the centers of the primary and secondary coils, centimeters (as in Fig. 6)
η = fractional deviation of a curve from a straight line tangent to the curve at some reference point
$S_m=\dfrac{dM}{dx}$ (inductance sensitivity)
μ = magnetic permeability
ϵ = dielectric constant
σ = conductivity
L = length
f = frequency $\left(f=\dfrac{\omega}{2\pi}\right)$
C_1 = constant for specific parameter values
C_2 = constant for specific parameter values

Analysis

Phenomenologically a mutual-inductance transducer of the eddy-current perfect-conductor type operates as follows. The voltage induced in the open-circuit secondary coil by the primary coil current is diminished by bringing a metal plate near the secondary coil; the closer the plate, the greater the decrease in induced voltage. This induced voltage is proportional to the effective mutual inductance between the primary and secondary coils, and can be calibrated as a function of the separation between the coil system and the reference plate.

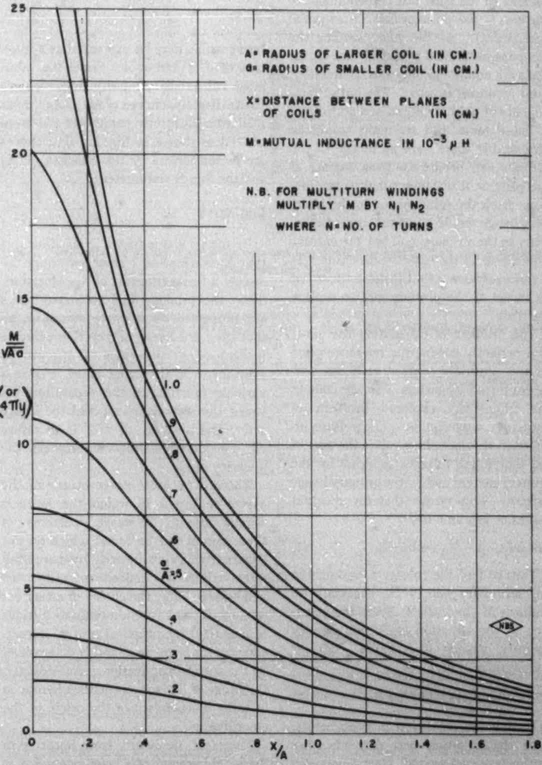
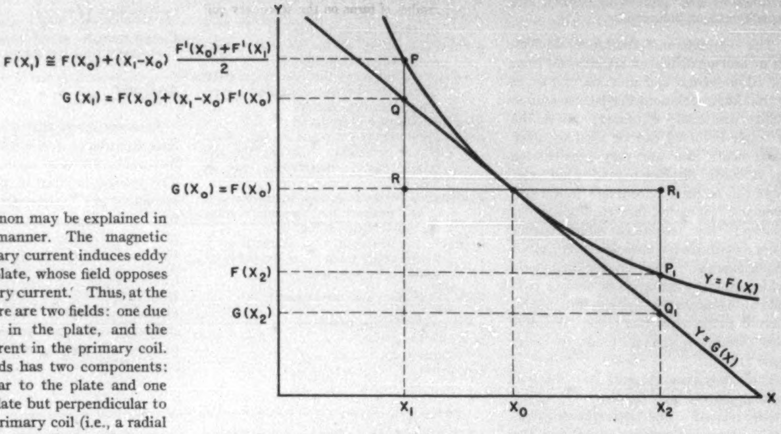


Fig. 3. Mutual inductance between two single-turn coaxial coils versus displacement between them

Fig. 4. Reference figure for definition of range of linearity


$F(X_1) \cong F(X_0)+(X_1-X_0)\dfrac{F'(X_0)+F'(X_1)}{2}$
$G(X_1) = F(X_0)+(X_1-X_0)F'(X_0)$
$G(X_0) = F(X_0)$
$F(X_2)$
$G(X_2)$
Y
P
Q
R
R_1
P_1
Q_1
$Y = F(X)$
$Y = G(X)$
X_1
X_0
X_2
X

This phenomenon may be explained in the following manner. The magnetic field of the primary current induces eddy currents in the plate, whose field opposes that of the primary current. Thus, at the plate surface there are two fields: one due to the currents in the plate, and the other to the current in the primary coil. Each of the fields has two components: one perpendicular to the plate and one parallel to the plate but perpendicular to the axis of the primary coil (i.e., a radial component). At the plate surface the components of field perpendicular to the plate are equal and opposite in direction; their resultant is zero. The radial components of field are equal and additive.

These same field strengths would be produced if there were an image[1] of the primary coil behind the front surface of the plate at a distance equal to the distance from the primary coil to the plate. This image would be of the same dimensions as the primary coil but the current would flow in the opposite direction. It is this equivalence that justifies the use of an image of the primary coil to replace the plate.

The problem of calculating the effect of a perfectly conducting reference plate on two mutually coupled coils can now be converted by means of image coils to the extensively explored problem of mutually coupled coils. The voltage induced in the secondary coil is the sum of the individual voltages induced by the primary current and by the primary image current. This means that for constant values of M_{12} and M_{12}'

$$E_2 = j\omega I_1(M_{12}-M_{12}') = E_{12}-E_{12}' \qquad (1)$$

Thus to find the voltage E_2 induced in the secondary coil, it is necessary to evaluate M_{12} and M_{12}'. From the geometry of the situation and from electromagnetic theory, it is found[2] that for multiturn windings

$$M = \sqrt{Aa}\,n_1 n_2 4\pi y \times 10^{-9} \text{ henrys} \qquad (2)$$

(A and a are measured in centimeters, since the equation was derived using centimeter-gram-second electromagnetic units.)

Values of $4\pi y$ have been tabulated[3] in terms which may be expressed as a function of a/A and x/A. Since the tabulated values are somewhat inconvenient to use directly, curves of $M/\sqrt{Aa}$ versus x/A with the family parameter a/A were plotted as shown in Fig. 3. The curves are dimensionless in the abscissa (x/A) and the family parameter (a/A).

LINEARITY

In many applications, it is desirable that E_2 in equation 1 be at least approximately a linear function of displacement. Since the primary and secondary coils are supported on a probe structure, so that they are fixed in position relative to each other, the value of M_{12} will be constant. On the other hand, M_{12}' is a variable function of the separation between the secondary coil and the image coil. The linearity of M_{12}' is therefore the consideration which determines the linearity of E_2.

There is no finite range of any of the curves in Fig. 3 in which the curve is strictly linear. However, portions of these curves can be found which do not depart from strict linearity by more than some arbitrarily defined quantity. As will be seen later, the choice of a zero (or reference) point (used in defining a range of linearity) for practical mutual-inductance curves is optional and is determined by the specific application. Accordingly, the range of linearity is defined herein so as to be independent of the origin of the co-ordinates.

Therefore, departure from linearity in the neighborhood of a reference point x_0 (see Fig. 4) will be defined in terms of the ratio:

$$\eta = \left|\frac{\overline{PQ}}{\overline{QR}}\right| = \frac{|\Delta F(x)-\Delta G(x)|}{|\Delta G(x)|} \qquad (3)$$

where $F(x)$ is the curve and $G(x)$ the equation of a straight line tangent at x_0. Then a "range of linearity" is a range of values of x throughout which η remains less than a preassigned value.

$$F(x) \doteq F(x_0)+(x-x_0)\frac{F'(x_0)+F'(x)}{2} \qquad (4)$$

Also

$$G(x) = F(x_0)+(x-x_0)F'(x_0) \qquad (5)$$

By substituting values of $F(x)$ and $G(x)$ from equations 4 and 5, the desired range of linearity from $x=x_1$ to $x=x_2$ corresponds to x_1 in the equation

$$F'(x_1) = (1\pm 2\eta)F'(x_0) \qquad (6)$$

and to x_2 in the similar equation

$$F'(x_2) = (1\pm 2\eta)F'(x_0) \qquad (6A)$$

From curves of $F'(x)$ versus x (available in Fig. 5) it is possible to read the values of x_1 and x_2 corresponding to $F'(x_1)$ and $F'(x_2)$ computed from equations 6 and 6(A). The derivative curve in the neighborhood of $x=x_0$ will make it apparent whether the plus or minus sign should be used in equation 6 and 6(A). The curves in Fig. 5 make it apparent that the greatest range of linearity for any particular curve occurs when x_0 is chosen so that

$$F'(x_0) = \frac{F'(x_i)}{1+\eta} \qquad (6B)$$

where x_i = the value of x at the point of inflection. There are two such values of x_0, one on either side of the inflection point.

INDUCTANCE SENSITIVITY

These derivative curves are intimately related to sensitivity since

$$S_m=\frac{dM}{dx}=n_1n_2\sqrt{\frac{a}{A}}\frac{d\left(\frac{M}{\sqrt{Aa}}\right)}{d\left(\frac{x}{A}\right)}$$

$$=n_1n_2\sqrt{\frac{a}{A}}\text{ (ordinate of Fig. 5)} \qquad (7)$$

The maximum obtainable inductance sensitivity is not a function of A, so long as the ratio a/A is held constant. Also, for given values of x and a/A there is a value of A which will give greatest sensitivity. Table III shows ranges of linearity average sensitivity for $\eta=0.05$. It is apparent in Table III that increasing the range of linearity by reducing the ratio a/A entails a concomitant loss of sensitivity.

The effect of the conducting plate upon a coil system can be determined by substituting an image coil and current for the plate, and calculating the result with the aid of the equations and curves indicated in the foregoing. This situation is represented in Fig. 6 for zero current in the secondary coil. A typical gauge curve that results from calculating the effect of variable position of the metal plate is shown in Fig. 7.

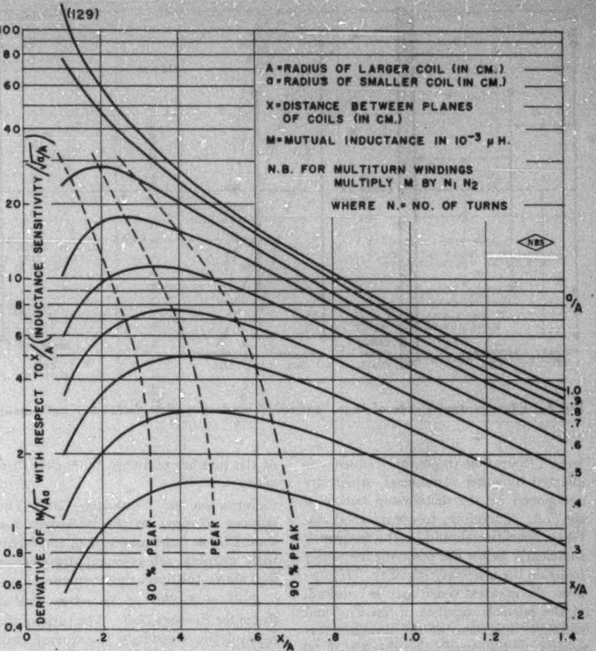


Fig. 5. Derivative of mutual inductance between two single-turn coaxial coils versus displacement between them

Design Recommendations

In view of the foregoing, certain design recommendations for obtaining optimum combinations of stability, sensitivity, and linearity are made.

Conductivity of the Plate: The material of the plate should have as high a conductivity as practicable. The plate will then more nearly simulate a perfect conductor, and the design curves (Figs. 3 and 5) will apply with a greater degree of reliability.

Relative Size of Primary and Secondary Coils: The primary coil should be the larger in order to provide greater heat-radiating surface, when the coils differ in size.

Position of Primary and Secondary Coils: From the point of view of sensitivity, the primary and secondary coils should be coplanar. However, if they are not coplanar, the primary coil should be the one farther from the plate so that the plate will have less effect upon the primary current.

Excitation Frequency: The excitation frequency should be as high as is consistent with avoiding the instabilities of resonance since the output increases with higher operating frequency. Also, a high excitation frequency decreases the depth of eddy-current penetration; hence the plate more nearly simulates a perfect conductor (zero depth of penetration).

In using as high a frequency as possible, however, the adverse effects of the unavoidable self-capacitance of the coils and their leads must be considered. Operation near the lowest self-resonant frequency is to be avoided because the output voltage is then a critical function of the exciting frequency and of the effective Q of the system. The Q itself is a critical function of the material, temperature, and mechanical dimensions of both the system and the surrounding ob-

Table III. Ranges of Linearity (Within 5 Per Cent) of the Static Mutual-Inductance Curves (Fig. 3)

	Range of Linearity				Average Inductance Sensitivity in this Linear Range (Equation 7) $\frac{\Delta M}{\Delta x}=\frac{\Delta M/\sqrt{Aa}}{\Delta x/A}\sqrt{a/A}$ (in 10^{-9} Henrys per Centimeter)
a/A	x_1/A	x_2/A	Extent of Linear Range $\Delta(x/A)$	Range of $M/\sqrt{Aa}$ (in 10^{-9} Henrys per Centimeter)	
0.8	0.12	0.29	0.17	18.2–13.8	23.1
0.7	0.17	0.39	0.22	13.3– 9.5	14.5
0.6	0.22	0.49	0.27	9.5– 6.5	8.6
0.5	0.26	0.55	0.29	6.7– 4.6	5.1
0.4	0.29	0.60	0.31	4.6– 3.1	3.06
0.3	0.31	0.64	0.33	2.9– 1.9	1.66
0.2	0.33	0.68	0.35	1.5– 1.0	0.64

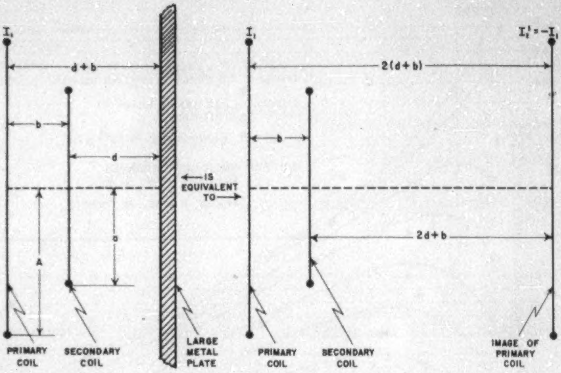


Fig. 6. Effective equivalence of large "perfectly" conducting plate (secondary coil open)

jects. Operation above this frequency is affected by other resonances, which are occasioned by the distributed nature of the coil. Therefore, to obtain good stability, operation should be at a frequency sufficiently below the resonant frequency to attain the necessary stability. If however, the greatest sensitivity is required, at the possible sacrifice of linearity and stability, operation at the resonant frequency is indicated. Sensitivity of 30 to 40 times the values calculated without considering resonance, have been obtained in this manner.

The actual value of self-capacitance is not easily predicted theoretically, since it depends upon the dielectric constants of the coil form and the wire insulation, and upon the specific manner in which the coil is wound. An indication of the value of this capacitance may be obtained from references 3 and 4.

Shielding of Leads: The leads to the primary coil should either be shielded from those of the secondary coil or be relatively fixed in position with respect to them. Undesired change in coupling would tend to mask the desired change in effective mutual inductance. It should be noted that the capacitance due to the shield will lower the self-resonant frequency.

Ampere-Turns of the Primary Coil: The ampere-turns of the primary coil should be as large as possible, since the output is proportional to this product. It is readily shown[5,6] that for a given allowable energy dissipation, coil volume and shape, frequency, material, etc., the maximum value of ampere-turns does not depend upon either the current or the number of turns, but upon their product.

Turns on the Secondary Coil: The number of turns on the secondary coil should be as large as possible, consistent with keeping self-resonant frequencies well above the operating frequency.

Ratio of Radius of Secondary Coil to that of the Primary Coil: The ratio of the radii of the secondary to primary coils (a/A) should be close to unity for relatively high sensitivity. Reference to Table III and Fig. 5 indicates this directly. As the sensitivity increases because of an increase in the ratio a/A, there is a corresponding but relatively small decrease in the range of linearity and inward change of its position.

Size of Primary Coil: The range of linearity increases as the radius of the primary coil increases for a constant value of a/A. This is the principal method of extending the range of linearity. Also, a larger primary coil has better heat-dissipating properties than a small one; however, the power requirement will be larger.

A larger primary coil radius may be accompanied by increased or decreased output voltage sensitivity (voltage change per unit displacement) depending upon what other changes in dimensions take place with the increase in coil radius. The analysis in Appendix I shows that if all linear dimensions of the probe system are increased by a scale factor h, then

$$S_e \doteq h^{\frac{1}{2}} S_e \tag{8}$$

where S_e is the voltage sensitivity of the scaled system. On the other hand, if A, a, and x are increased by a factor h, but the wire size and coil cross sections are kept unchanged, then

$$S_e' \doteq \frac{1}{h^{\frac{1}{2}}} S_e \tag{9}$$

(This particular result applies when the skin depth is small compared to the wire radius.) Thus in the one case voltage sensitivity is increased; in the other, decreased.

A probe with a large primary coil radius can be made to have a very high sensitivity. The curves shown in Fig. 5 indicate that the inductance sensitivity increases without limit as $(a/A) \rightarrow 1$ and $x/A(=2d+b/A) \rightarrow 0$. For a large A, a/A can be made to approach unity since the structural limitations on reaching this limit such as insulation thickness, minimum wire size, etc., are nearly constant, and hence represent smaller proportions of A. For the same reasons x/A can be made to approach zero.

Despite these considerations and possibly those of convenience and feasibility, the maximum coil size is limited by the

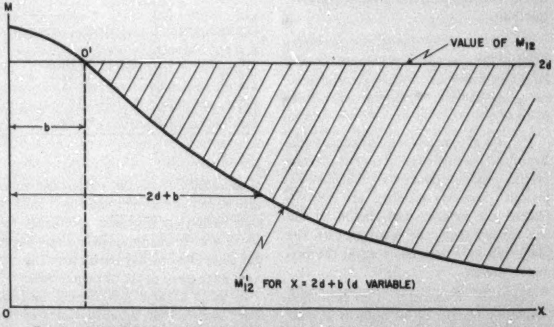


Fig. 7. Mutual-inductance curve showing effect of large perfectly conducting plate

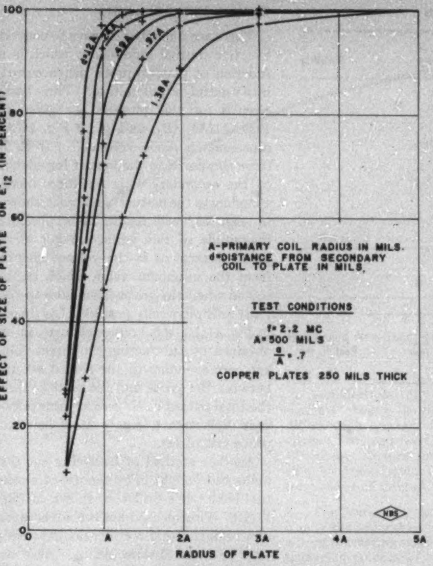


Fig. 8. Effect of finite plate size

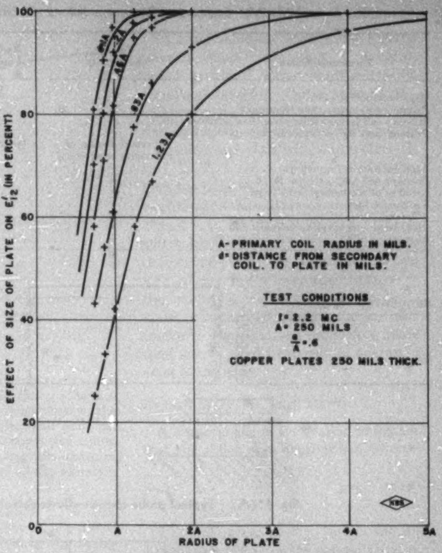


Fig. 9. Effect of finite plate size

extent of the plate. Theoretical considerations indicate that an actual plate has somewhat less effect than that calculated for an infinite plate. For a fixed percentage reduction in effect, the size plate required is a function of the distance between the coil system and the plate. Inserting the coefficient Γ in equation 1 to represent the effect of a finite plate

$$E = E_{12} - \Gamma E_{12}' \qquad (10)$$

represents the output voltage from a probe in the proximity of a perfectly conducting plate of limited size.

Theoretical calculations have not been carried to completion, but the results of two experimental cases giving the change in output due to a plate of limited size as a percentage of the change in output due to a plate six times the size of the larger (primary) coil are shown in Figs. 8 and 9 for different distances between the coils and the plate. These values apply for coplanar coils whose ratios, a/A, are 0.6 and 0.7 and do not necessarily apply for coils where a/A is widely different from these values, or where the coils are not coplanar. It appears that, in general, the greater the separation between coils and plate, the larger the plate necessary for a given effect.

Regulation of $I\omega$: The product $I\omega$ of the exciting frequency and the exciting current should be regulated to a constant value for stability of output. This follows directly from equation 1. Regulation of this product involves little difficulty. It can be attained by maintaining a constant (open-circuit) voltage across the secondary winding of an auxiliary transformer whose primary is excited by the current I_1. This is shown in Fig. 10.

General Discussion of Recommendations: Obviously, all of the recommendations cannot be satisfied simultaneously. Compromises must be made in accordance with the results desired. Table IV summarizes the effect of the recommendations on voltage sensitivity, range of linearity, and stability. A tendency to increase is designated by a (+) sign and a tendency to decrease is indicated by a (−) sign.

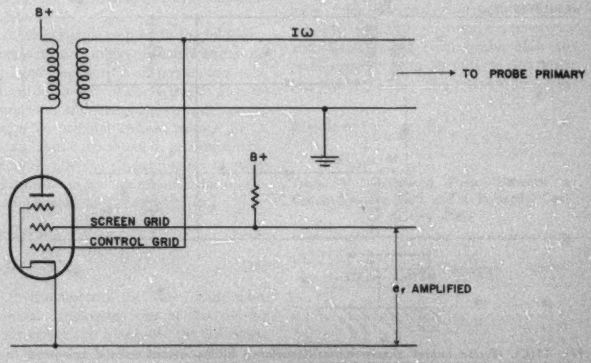


Fig. 10. Basic regulating circuit

Table IV. Summary of Recommendations and Effects

Recommendation	Voltage Sensitivity	Range of Linearity	Stability
High conductive plate	+		
Primary coil larger than secondary			+
Primary coil farther from plate than secondary coil			+
Primary and secondary coils coplanar	+ (beyond region of maximum sensitivity)	+ (generally)	
High excitation frequency	+		−
Primary coil ampere-turns large	+		
Number of secondary turns large	+		−
Ratio a/A high	+	−	
A large	±	+	
Plate large compared to primary coil	+		
Regulation of $I\omega$	+		+

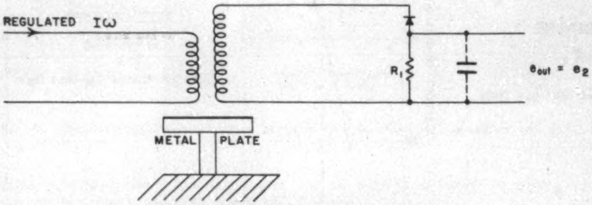


Fig. 11(A). Typical probe circuit—diode detector

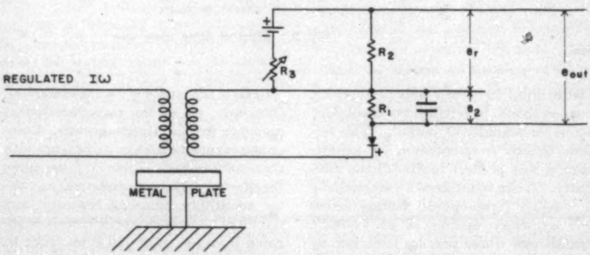


Fig. 11(B). Probe output circuit—diode detector with variable reference for output voltage

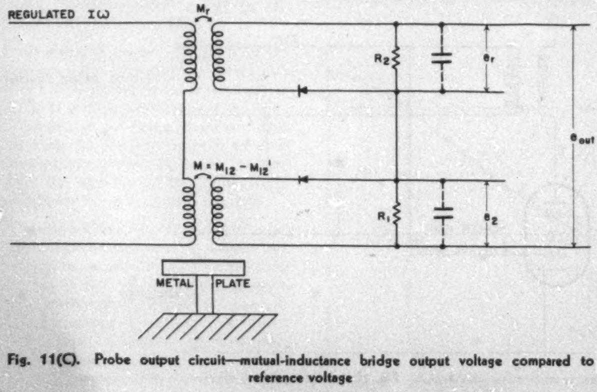


Fig. 11(C). Probe output circuit—mutual-inductance bridge output voltage compared to reference voltage

TYPICAL PROBE CIRCUITS

There are a number of ways to convert effective mutual inductance, which is a function of plate-to-probe displacement, into a useful output voltage. Three basic circuits for this purpose are given in Figs. 11(A), (B), and (C). Fig. 11(A) represents a diode voltmeter. If R_1 is large compared to the output impedance of the secondary coil, a voltage corresponding to the mutual inductance, shown in Fig. 12, will result. The absolute magnitude of this voltage is not of as much interest as is the voltage change from the minimum value which is obtained when the probe is as close to the plate as is physically possible. As shown in Fig. 11(B), this voltage change can be obtained by subtracting the direct voltage corresponding to the closest setting between the probe and the plate so that the total output e_{out} is zero at some reference diplacement (e.g., contact between probe and plate).

Another method of balancing out the undesired voltage is by the use of a mutual-inductance bridge as shown in Fig. 11(C). This method has the advantages of a bridge in that rectifier nonlinearities may be offset to some extent. Also the mutual inductance of the comparison coils may be changed to produce a balance.

The voltage output e_r from the comparison or reference coil of fixed mutual-inductance M_r may be used to control the value of the product $I\omega$ by appropriate feedback. Such a scheme is indicated in Fig. 10. Here, the voltage on the screen grid of a pentode oscillator is varied in accordance with the variations of the reference output voltage e_r. Complete circuits, showing regulation and buckout schemes as described, are given in references 7, 8, and 9.

FORCE ON PLATE DUE TO THE PRIMARY CURRENT

In some applications it is necessary to consider the effect of electromechanical forces on the plate resulting from the presence of the probe. As will be seen, these forces are rather small. There will be no effect because of the presence of the secondary coil since there is no current through this coil.

The force on the primary coil resulting from the presence of a large perfectly conducting plate is the same as that resulting from the image coil. Since "action equals reaction," this force is the same in magnitude as the total force on the plate. Since the current in the image coil is opposite in direction to that of the primary coil, the force is one of repulsion.

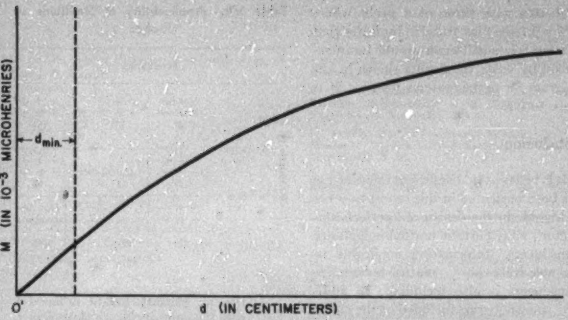


Fig. 12. Effective mutual inductance versus distance d from secondary coil to plate

An equation for evaluating the force[5] between a coil and its image is

$$F=-n_1^2 i_1^2 \frac{dM}{dx} 10^{-11} \tag{11}$$

Here x is the distances between the primary coil and its image, or twice the distance from the primary coil to the plate. If the axis of the primary coil is perpendicular to the plate, dM/dx becomes the ordinate shown in Fig. 5 for the curve $(a/A)=1$.

Several typical values of force corresponding to values of i_1, n_1, and x/A are given in Table V to show the order of magnitude of the force involved. This equation applies to the case of a perfectly conducting plate, infinite in extent. The size limitations of a practical plate will decrease the force below this value; finite resistivity of the plate will also reduce the force.

GAUGES FOR DYNAMIC MEASUREMENTS

In the previous sections changes in mutual inductance were considered as a function of a static configuration. If there is a relative motion, the probe output will be a function of the rate of change of mutual inductance with respect to time (dM/dt) as well as of the rate of change of current with respect to time (dI/dt). That is

$$e=\frac{d(MI)}{dt}=M\frac{dI}{dt}+I\frac{dM}{dt} \tag{12}$$

An analysis of the effect of this motion shows that the effect is similar (but not identical) to amplitude modulation. There is no error in using static considerations if the following conditions are met:

1. The frequency of the exciting current is higher than the highest frequency component of the mechanical vibration.
2. The bandwidth of the circuits which pass the carrier or exciting frequency and its modifications due to the mechanical motions should be flat from below the difference frequency of electrical and mechanical frequencies to above their sum. Of course, circuits following the detector (if any) should be flat over all the vibration frequencies of interest.
3. The equivalent amplitude modulation should not be increased beyond 100 per cent when subtracting carrier voltage by bridge balancing processes.

Application of Electrodynamic Similitude Principles to Gauge Development

In the preceding sections no methods were indicated for predicting the effects of reference plates of types different from extensive perfect conductors. No general theory which will enable prediction of these effects has been found in this investigation. However, it is often possible to obtain the desired prediction from tests on an appropriate scaled model by use of the principle of electrodynamic similitude.

That principle as used in the following means that in a given electric circuit or coil configuration, if all parameters of a given dimension (such as length or conductivity) are changed by a scale factor, then the original output voltage for a given input current can be maintained by properly altering all parameters of other dimensions in accordance with the dimensional equations 13 and 14[10]

$$\mu\epsilon(fL)^2=C_1 \tag{13}$$

$$\mu\sigma fL^2=C_2 \tag{14}$$

In the material to follow, while more general conclusions can be derived by considering μ as a variable, the discussion will be confined to nonferromagnetic materials where the μ is constant.

SCALING FROM EXPERIMENTAL RESULTS

In a mutual-inductance probe, a given primary current will result in a corresponding voltage output which will be the same both before and after scale modifications on the probe. If either the design or experimental curves for a model transducer and plate with parameters A, a, d, b, R, t, σ, and f are available, then the same magnitudes of E and I will be obtained if the scaled system has parameters hA, ha, hd, hb, hR, ht, σ/h, and f/h, where h is the scaling factor, R the plate radius, and t the plate thickness. When $h>1$, a larger probe is required to increase the operating distances, and the plate must be correspondingly larger and thicker. The electrical quantities, plate conductivity, and frequency of operation must be reduced.

EFFECT OF FINITE SKIN DEPTH

A consequence of the scaling process is that the scaled skin depth is h, δ, where[11]

$$\delta=\sqrt{\frac{C_2}{\mu\sigma f}} \tag{15}$$

The scaling was used to justify the dimensionless form for the presentation of the experimental measurements of the effects of limited plate size in Figs. 8 and 9.

Another example of the conclusions that may be obtained by using the principle of similitude is the following: An empirical equation was obtained from a series of measurements at different frequencies which showed how the skin depth (in mils) modified the output of a probe system. The measurements were made with a probe coil whose primary coil radius was 500 mils, secondary coil radius 350 mils (i.e., $a/A=0.7$), frequencies ranging from 100 kc to 1 megacycle, a large thick circular copper plate of resistivity (ρ) 2.16 microhm-centimeters, and displacement from probe coils to plate ranging from 50 mils to 850 mils ($0.1 \leq d/A \leq 1.7$). The relation found is

$$E_2=E_{12}-\frac{1}{1+0.0066\delta}E_{12}' \tag{16}$$

Table V. Calculated Force Between a Current-Carrying Coil and a Perfectly Conducting Plate

$\frac{x}{A}$	$\frac{d+b}{A}$	n	i, Amperes	F, Dynes
0.1	0.05	12	1	−185
0.1	0.1	12	1	− 95.5
0.4	0.2	12	1	− 40.0
1.0	0.5	12	1	− 10.8
2.0	1.0	12	1	− 2.18

1 ounce≐27,800 dynes.

Table VI. Percentage Reduction from Maximum Value of E_{12}' when $a/A=0.7$ and $0.1 \leq d/A \leq 1.7$

Frequency	Material	Skin Depth, Mils	Per-Cent Reduction in E_{12}'
100 kc	copper	9.2	5.7
100 kc	brass	16.8	10.0
1 megacycle	copper	2.91	1.9
1 megacycle	brass	5.3	3.4

An additional series of experiments with a corresponding brass plate ($\rho=7.17$ microhm-centimeters), using the same probe, range of frequencies, and range of displacements, was found to be a close approximation by the same relation 16.

Equation 16 indicates that the effective E_{12}' is less than E_{12} by a percentage depending upon δ in the experiments. This percentage reduction is tabulated in Table VI for copper and brass at the two extreme frequencies used in the experiments. It appears reasonable that for the given probe and a given frequency, the relation 16 applied over a continuous range of conductivities from that of copper to that of brass and gives a monotonic variation in the percent reduction in E_{12}' over this range of conductivities.

By application of similitude, any probe having $a/A=0.7$ operating at any frequency in the neighborhood of an extensive thick plate of any conductivity can be scaled to a probe whose primary coil radius is 500 mils, and to a corresponding frequency and plate conductivity. The converted frequency and conductivity will fall into one of four main groups:

1. Both fall within their respective limits in the experiments. Per-cent reduction of E_{12}' of original system can be calculated.
2. Only one falls within the range of the experimental limits. An upper or lower limit on the per-cent reduction of E_{12}' can be calculated from the experimental data referred to in item 1 as indicated in Table VII.
3. Both fall outside the range of experimental limits in a manner such that each tends to have an effect in the same direction on the per-cent reduction in E_{12}'. Per-cent reduction in E_{12}' is less than 1.9 per cent or greater than 10 per cent.
4. Both fall outside the range of experimental limits in a manner such that each tends to have an effect in opposite directions on the per-cent reduction in E_{12}'. The per-cent reduction in E_{12}' cannot be predicted without additional information.

The effect of the four groups is shown in Table VII.

Although the experiments and resulting inferences apply to probes with a ratio $a/A=0.7$, a comparison of the experimental results with a probe where $a/A=0.6$ with those of a probe where $a/A=0.7$(see Figs. 8 and 9) indicates that a relatively small error would be introduced by using the results shown in the diagram for probe systems whose a/A is close to 0.7.

Conclusion

It is believed that sufficient information has been presented in this report to serve as a guide in the design of "perfect-conductor" eddy-current mutual-inductance transducers. Information applicable to movable-coil-type mutual-inductance transducers is also included. In addition, some notes are included on the modifications that will occur when resistive plates are used. An indication is given of how to determine when the errors, made by applying perfect-conductor results to practical resistive conductors, are small.

Although a method of scaling experimental models is mentioned, the information on eddy-current mutual-inductance transducers when used with resistive plates is incomplete. Also, no information on ferromagnetic plates is given. The theoretical prediction of results for these cases does not appear to be available. The solution of this problem would be of value in many problems that require the use of resistive conductors. These include the determination of the position of a thin diaphragm, or of the top of a column of mercury, problems in the field of induction heating, and the measurement of the electrical properties of alloys for such purposes as inspection, the determination of hardness, or the thickness of plated metal film.

Appendix I. Voltage Sensitivity Versus Primary Coil Radius

The voltage sensitivity is affected differently by an increase in primary coil radius depending upon whether or not the wire size and coil cross sections are correspondingly increased. In both cases to be discussed, it will be assumed that A, a, b, and d of the probe configuration are multiplied by a factor h (greater than unity).

The following symbols and associated meanings will be used:

I = maximum permissible current in the primary coil for a fixed temperature rise
I' = maximum permissible current in the primary coil of radius hA for the same temperature rise
ω = maximum angular frequency of excitation for primary coil of radius consistent with avoiding resonance effects
ω' = maximum angular frequency of excitation for primary coil of radius hA with the same avoidance of resonance effects
α = ratio of a-c resistance to d-c resistance in the primary coil
P = power dissipated in primary coil
A_r = heat-radiating area of primary coil
R = total d-c resistance of primary coil
s = total length of wire in the primary coil
d = diameter of wire of the primary coil
S_e = voltage sensitivity
S_e', P', A_r', R', s', d', and α' will have corresponding meanings for the primary coil when modified, as indicated.

Table VII. Applicability of Similitude to Probes

Conductivity	Frequency		
	Not determined (4)	+ + + + + + .N (2) .N′	+ + + + + + (3)
σ copper			
	− − − − − − (2)	Calculable (1)	+ + + + + + (2)
σ brass			
	− − − − − − − − − (3)	− − − − − − − − − (2)	Not determined (4)
		100 kc	1 mc

As a first case, assume all linear dimensions of the coil system are multiplied by a factor h. A fixed temperature rise means that

$$\frac{P'}{A_r'}=\frac{P}{A_r} \tag{17}$$

Substituting resistive losses for P's

$$I'=h^{(3/2)}I\sqrt{\frac{\alpha}{\alpha'}} \tag{18}$$

Since the self-inductance and self-capacitance of a coil each vary with the average radius of the coil[13]

$$L'C'=h^2LC \tag{19}$$

Hence

$$\omega'=\frac{\omega}{h} \tag{20}$$

From equation 1

$$S_e=I\omega\frac{dM}{dx} \tag{21}$$

and, as dM/dx is independent of A for a given a/A and x/A as previously noted, S_e is proportional to $I\omega$. Therefore, from equations 18 and 20

$$S_e'=h^{1/2}S_e\sqrt{\frac{\alpha}{\alpha'}}=h^{1/2}S_e \tag{22}$$

Since all linear dimensions and frequency have been scaled, $\alpha'=\alpha$ (reference 13); thus the voltage sensitivity rises as A increases (under these assumed conditions).

As a second case, assume the wire size and coil cross sections remain unchanged while A, a, b, and d are increased by a factor h, as in the first case. Equation 19 is approximately true for this case. Then

$$I' \doteq I\sqrt{\frac{\alpha}{\alpha'}} \tag{23}$$

since the power dissipated per unit area of coil is constant.

Now[13]

$$\omega' \doteq \frac{\omega}{h} \tag{24}$$

Then

$$S_e' \doteq \frac{S_e}{h}\sqrt{\frac{\alpha}{\alpha}} = \frac{S_e}{h^{\frac{1}{2}}} \tag{25}$$

since case $\alpha' \doteq (1/h)\alpha$, when $\delta<<A$ (reference 13). Thus the voltage sensitivity drops as A increases.

References

1. ELECTRICITY AND MAGNETISM, **J. C. Maxwell.** Clarendon Press, Oxford, England, vol. 1, chap. XI, 1892.

2. TABLES FOR FACILITATING THE CALCULATION OF SELF-INDUCTANCE OF A CIRCULAR COIL AND OF MUTUAL INDUCTANCE OF COAXIAL CIRCULAR CURRENTS, **H. Nagaoka, S. Sakurai.** *Scientific Papers of the Institute of Physical and Chemical Research*, Tokyo, Japan, Sept. 1927, table 2.

3. RADIO ENGINEERS' HANDBOOK (book), **F. E. Terman.** McGraw-Hill Book Company, Inc., New York, N. Y., 1943, p. 101.

4. DISTRIBUTED CAPACITANCE CHART, **P. H. Massaut.** *Electronics*, New York, N. Y., vol. 11, March 1938, p. 31.

5. ELECTROMAGNETIC DEVICES (book), **H. C. Roters.** John Wiley and Sons, Inc., New York, N. Y., 1941.

6. HEAT TRANSMISSION (book), **W. H. McAdams.** McGraw-Hill Book Company, Inc., New York, N. Y., 1942.

7. TECHNICAL DATA ON ELECTRONIC MICROMETER. *Electronics*, New York, N. Y., vol. 20, Nov. 1947, p. 172.

8. SENSITIVE MERCURY-LEVEL DETECTING UNIT FOR MANOMETERS, **W. E. Williams, Jr.** *Journal of Research*, National Bureau of Standards, Washington, D. C., vol. 48, Jan. 1952, p. 54.

9. AN ELECTRONIC CIRCUIT FOR MEASURING THE DISPLACEMENT OF PRESSURE-SENSITIVE DIAPHRAGMS, **M. L. Greenough, W. E. Williams, Jr.** *Ibid.*, vol. 46, Jan. 1951, p. 5.

10. ELECTROMAGNETIC THEORY (book), **J. A. Stratton.** McGraw-Hill Book Company, Inc., New York, N. Y., 1941, p. 488.

11. **Terman,** *op. cit.*, p. 34.

12. RADIO INSTRUMENTS AND MEASUREMENTS. *Circular C 74*, National Bureau of Standards, Washington, D. C., 1937, sections 64 and 67.

13. **Terman,** *op. cit.*, p. 31.

14. PRINCIPLES OF ELECTRICITY AND MAGNETISM (book), **L. Page, N. I. Adams, Jr.** D. Van Nostrand Company, Inc., New York, N. Y., 2nd ed., 1949, p. 364.

15. OIL-FILM THICKNESS INDICATOR FOR JOURNAL BEARINGS, **M. L. Greenough.** *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 589–95.

16. ELECTRONIC STEPMETER REVEALS MECHANICS OF WALKING. *Technical News Bulletin*, National Bureau of Standards, Washington, D. C., vol. 35, no. 4, April 1951.

17. WIDE RANGE CALIBRATOR FOR VIBRATION PICKUPS, **W. A. Yates, M. Davidson.** *Electronics*, New York, N. Y., vol. 26, Sept. 1953, p. 183.

L. 74
1
55

A Fast Response Magnetic Servo Amplifier

J. W. KALLANDER
ASSOCIATE MEMBER AIEE

IN SERVOMECHANISMS used in military applications, one of the weakest links is the electronic amplifier. The electronic amplifier is undesirable primarily because of its fragility, maintenance requirements, low power efficiency, and often objectionable warm-up time. The magnetic amplifier is much more rugged, has a much longer life, has a greater power efficiency, requires much less maintenance, and has no warm-up time. Thus, one of the most obvious and potentially fruitful applications of magnetic amplifiers is in the control of 2-phase motors in servomechanisms.

Although the magnetic amplifier field made tremendous strides after World War II, the inherently long time delay of the magnetic amplifier seriously limited its applications to servomechanisms until the invention of the "half-cycle response time" magnetic amplifier[1,2] several years ago. This magnetic amplifier has a time delay of from 1/2 to 1 cycle of the supply frequency for full output, i.e., 0.0083 to 0.0167 second for a 60-cycle amplifier or 0.0012 to 0.0025 second for a 400-cycle amplifier, which is sufficient for many servo applications.

Since that time only one half-cycle response time magnetic servo amplifier[3] has been described in the literature. Consequently, it is believed that additional magnetic servo amplifiers of the half-cycle type will find useful applications. Thus, it is the purpose of this report to describe the theory and operation of a new magnetic servo amplifier.

The following major conditions have been used as a guide in deducing these circuits:

1. The output of the magnetic amplifier must contain a phase-reversible a-c component of the fundamental frequency capable of controlling a 2-phase motor.

2. The input circuit must be sensitive to a phase-reversible a-c signal of the fundamental frequency or to a reversible-polarity d-c signal.

3. The amplifier must have a fast response time, preferably 1/2 cycle.

4. The amplifier must be capable of driving the servomotor and load, i.e., must meet their power requirements and must be able to operate into the back electromotive force of the motor.

5. The amplifier should be capable of presenting a high impedance to the input signal source.

Nomenclature

e_{ac} = gating-circuit supply voltage
e_{ac}' = control-circuit supply voltage
e_c = control voltage
R_c = control-circuit resistor
ϕ_c = control phase of 2-phase servomotor
N = ratio of gating-winding turns to reset (control)-winding turns
N_g = number of gating-winding turns
N_c = number of reset (control)-winding turns
ωt = radians
ω = angular velocity, radians per second
B = flux density
Ni = magnetomotive force
k = ratio of control-circuit supply voltage e_{ac}' to gating-circuit supply voltage referred to the control circuit e_{ac}/N
α = quiescent firing angle
e = reset (control)-winding voltage
e_{I} = core I reset (control)-winding voltage
e_{II} = core II reset (control)-winding voltage
e_R = control-resistor voltage
$e_{R\mathrm{I}}$ = core I control-resistor voltage
$e_{R\mathrm{II}}$ = core II control-resistor voltage
I_m = maximum magnitude of the reset-circuit magnetizing current
E_R = product of the maximum magnitude of the reset-circuit magnetizing current I_m and the control resistor R_c
L = magnitude of the maximum volt-second integral, referenced to the reset (control)-winding, available for application to a core during its gating half-cycle (ghc)
Δe = change in the reset (control)-winding voltage caused by the control voltage
Δ = integral of Δe over the reset half-cycle (rhc)
i_m = control-circuit magnetizing current

Description of Circuits

All magnetic cores used in the circuits described in this paper are of the so-called

---

Paper 55-57, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted June 14, 1954; made available for printing November 29, 1954.

J. W. KALLANDER is with the Naval Research Laboratory, Washington, D. C.

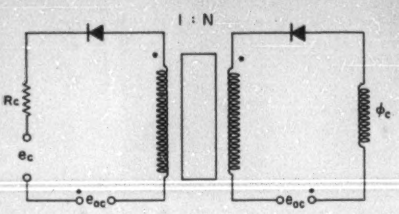


Fig. 1. Ramey single-core magnetic amplifier

"square loop" core material. The convention of indicating square-loop core material by means of a rectangle drawn between the windings on a core has been adopted in the circuit diagrams. The ends of the windings on a core having the same instantaneous polarity are dotted. This convention applies to the cores individually and does not indicate their relative polarities. The ends of transformer windings having the same instantaneous polarity are also dotted. In this case, the dotted ends of all transformer windings in a circuit have the same instantaneous polarity.

RAMEY SINGLE-CORE MAGNETIC AMPLIFIER

The Ramey single-core magnetic amplifier,[1,3] Fig. 1, is a very simple magnetic amplifier which meets the response-time and output rating conditions of the preceding section. One single-core amplifier is sensitive to an a-c control signal within a given 180-degree phase range or to a d-c control signal of a given polarity, yielding an output which will drive a 2-phase servomotor in a given direction; in the same circuit, a control signal of the opposite 180-degree phase range or polarity yields no output. A second single-core amplifier can be made sensitive to a control signal of the opposite 180-degree phase range or of the opposite polarity, yielding an output of the phase which will drive the 2-phase motor in the direction opposite to that of the first single core amplifier. Again, the second single-core amplifier does not respond to a control signal of the 180-degree phase range or polarity opposite to that for which it was designed. Thus, a combination of two single-core magnetic amplifier should also satisfy the reversible input and output conditions, forming a magnetic servo amplifier which meets the first four conditions imposed in the foregoing.

Two single-core amplifiers, each with half-wave output, can be combined to fire on alternate half-cycles, thus giving a unidirectional output.[4] A Fourier analysis of such an output wave shows no fundamental frequency component when both cores are firing at the same angle during their respective gating half-cycles. Thus, the amplifier can be biased to give a quiescent output current in which the outputs on consecutive half-cycles are equal.

The output-voltage wave shape of such an amplifier biased to fire at 90 degrees in its quiescent state is shown in Fig. 2(A). The solid lines give the amplifier output voltage; the dashed lines indicate the supply voltage rectified. As already stated, the quiescent output has no fundamental frequency component; hence, no net output torque would be supplied by a 2-phase servomotor excited by such an output. Fig. 2(B) shows the output wave form for a control signal of a given 180-degree phase range or of a given polarity. The output of one core, say core I, has been advanced to a firing angle of β while the output of the other core, say core II, has been retarded to a firing angle of $(\pi-\beta)$ on its ghc. The output will now have a fundamental frequency component. Fig. 2(C) gives the output wave shape for a control signal of the opposite 180-degree phase range or of the opposite polarity. The output of core I is now retarded to a firing angle of β while the output of core II is advanced to a firing angle of $(\pi-\beta)$ on its ghc. The output, now has a fundamental frequency component of the opposite phase to that of Fig. 2(B).

A Fourier analysis of the output gives the ratio of the fundamental frequency output component to the supply voltage as

$$\frac{b_1}{E}=\frac{1}{\pi}\left(\frac{\pi}{2}-\beta+(1/2)\sin 2\beta\right) \qquad (1)$$

in phase with the supply voltage. Fig. 3 gives the variation of b_1/E with the firing angle β of the first core.

It is advantageous to bias an amplifier in this manner to decrease or nullify completely the variation in phase of the fundamental frequency component as its magnitude is varied, to furnish a quiescent output damping current, and to realize increased gain. In the case discussed, i.e., a quiescent firing angle of 90 degrees, the phase of the fundamental frequency

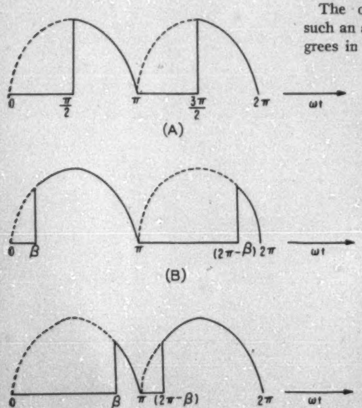


Fig. 2 (left). Output wave shapes of two single-core magnetic amplifiers operating in push-pull

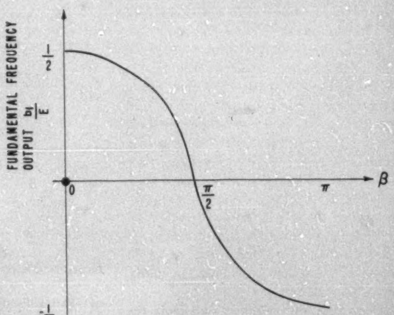


Fig. 3 (right). Output versus firing angle of two single-core magnetic amplifiers operating in push-pull

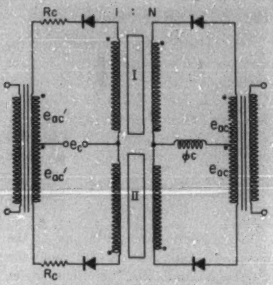


Fig. 4. New magnetic servo amplifier, a-c control

component of the output is constant and is in phase, or 180 degrees out of phase, with the supply voltage.

NEW MAGNETIC SERVO AMPLIFIER—A-C CONTROL

One means of combining two single-core amplifiers in a manner satisfactory for 2-phase servomotor operation is given in Fig. 4. This circuit consists essentially of two single-core amplifiers connected in parallel across a common signal source and a common load, as shown. The load circuit of the amplifier so obtained is the same as that of a magnetic servo amplifier described by W. Geyger.[4] Geyger[5] has recently described another magnetic servo amplifier in which the control circuit, rather than the load circuit, closely resembles the amplifier described here. In the Ramey single-core amplifier, $e_{ac}'=e_{ac}/N$ and the control resistor R_c is absent or small, furnishing a quiescent output current if desired. This circuit with the two single-core amplifiers unaltered is satisfactory for signal sources of low output impedance requiring a low impedance load. The circuit, however, is unsatisfactory for operation with a high impedance in the control circuit, such as required by a synchro control transformer, because the large impedance would absorb most, if not all, of the control-circuit supply voltage on the rhc, leaving little if any, voltage for the reset winding. Thus, the amplifier would only partially reset, if at all, on the rhc. This limitation can be removed by increasing the control-circuit supply voltage e_{ac}' from e_{ac}/N to ke_{ac}/N where $k>1$. In addition, on the ghc, when the control voltage is adding to the control-circuit winding voltage of the gating core, a control voltage of $e_{ac}'-e_{ac}/N$ is necessary before transformer action commences in this circuit. Thus, with $e_{ac}'=e_{ac}/N$, transformer action will take place for all e_c but, for $e_{ac}'=ke_{ac}/N$, transformer action will only occur when $e_c>(k-1)e_{ac}/N$. For no quiescent current

$$\int_{\text{rhc}} e_{ac}/N dt$$

must appear across the reset windings during their rhc. Thus, the value of k is determined from the following equation

$$\int_{\text{rhc}} i_m R_c dt=\int_{\text{rhc}}(k-1)e_{ac}/N dt,\ k\geq 1 \qquad (2)$$

where

i_m = the control-circuit magnetizing current
R_c = the control resistor determined by the minimum impedance required by the control source
N = the ratio of the gating-winding turns to the reset-winding turns
e_{ac} = the supply voltage in the gating circuit

If a quiescent current is desired, it can be obtained simply by increasing the value of the control resistors R_c above the value obtained in the preceding equation. Quantitatively, let α be the ratio of the desired average quiescent current on a half-cycle to the average of the maximum output current on a half-cycle, then the preceding equation for determining k can be modified to

$$\int_{\text{rhc}} i_m R_c dt=\int_{\text{rhc}}(k-1+\alpha)e_{ac}/N dt,$$
$$k\geq 1,\ 0\leq\alpha\leq 1 \qquad (3)$$

See equation 5 for an equivalent definition of α. When the control source has a high impedance R_s, R_c in the foregoing equations can be replaced by R_c+R_s.

It should be noted that the power dissipated in the control resistors is furnished by the control-circuit supply voltage. Thus, the voltage sensitivity can be made large, since the lower limit on the number of reset-winding turns is determined by practical considerations.

NEW MAGNETIC SERVO AMPLIFIER—D-C CONTROL

Fig. 4 can be changed into a d-c controlled circuit simply by reversing the sense in which one of the control circuits of the single-core amplifiers is paralleled across the signal source. This circuit is shown in Fig. 5. The foregoing symbols and equations are also valid for this circuit.

NEW MAGNETIC SERVO AMPLIFIER—IMPROVED GAIN

The characteristics of this circuit, when in operation, are such as to give effective hysteresis loops resembling those of Fig. 6 if perfectly rectangular hysteresis loops are assumed. Fig. 6(A) shows a typical quiescent-state loop with the core circling about half of its hysteresis loop. Figs. 6(B) and (C) show the steady-state hysteresis loops circled by the two cores of an amplifier when a given control signal is applied. The control signal hinders the reset of the core whose hysteresis loop is shown in Fig. 6(B), say core I, and aids the reset of the core whose hysteresis loop is shown in Fig. 6(C), say core II. The control-circuit current in Fig. 6(B) is decreased and that in Fig. 6(C) is increased. Thus, the core I control-resistor voltage drop will decrease, leaving a greater voltage drop across core I during reset than if the control-circuit current had remained constant. Thus, the control signal does not all appear across the reset winding but divides between the reset winding and the control resistor. Similarly, in core II, the control-voltage resistor drop increases so that the control signal divides between the reset winding and the control resistor instead of appearing solely across the reset winding. Thus, the gain of the new magnetic servo amplifier shown will decrease with increasing signal. This is because the control-circuit current and, hence, the control-resistor voltage drop on the reset half-cycles decrease as the core circles smaller and smaller hysteresis loops, and increase as the core circles larger and larger hysteresis loops.

The voltage drop across the control resistor can be prevented from decreasing appreciably by inserting a voltage source and rectifier across the control resistor, as shown in Fig. 7 for the a-c controlled

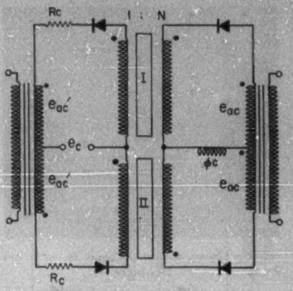
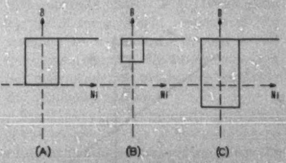


Fig. 5. New magnetic servo amplifier, d-c control



(A) (B) (C)

Fig. 6. Effective hysteresis loops of the new magnetic servo amplifier

ol. 74
. 1
955

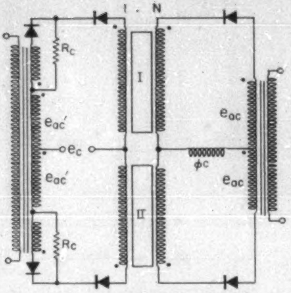


Fig. 7. New magnetic servo amplifier, a-c control, improved gain

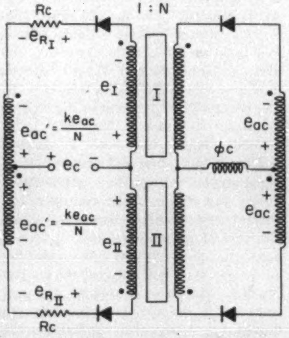


Fig. 8. New magnetic servo amplifier circuit for theoretical analysis

amplifier. They can be added to the d-c controlled amplifier in exactly the same manner with the same results. The voltage source is of the same frequency and phase as the supply voltages and of a magnitude equal to the magnitude of the quiescent-state control-resistor voltage drop. In the quiescent state, this voltage source furnishes no current but, as the magnetizing current decreases, tending to decrease the control-resistor voltage drop, this voltage source maintains the control-resistor current essentially constant, thereby maintaining the same voltage drop across the control resistor as in the quiescent state. Thus, on the rhc, most of the control-signal voltage appears across the reset winding.

Theoretical Analysis

The rectifier algebra of Schaefer[6] will be used to give a quantitative analysis of the new magnetic servo amplifier for an arbitrary signal-voltage wave form. The analysis will be based on the volt-second integral concept and depends on the fact that the output of the amplifier on any half-cycle depends solely on the signal voltage applied during the previous half-cycle. Perfectly square-loop cores, constant magnitude of the magnetizing current, and perfect rectifiers are assumed, and all resistance in the control circuit is included in the control resistors. The effect of the back electromotive force of the motor is not taken into account.

Fundamentally, rectifier algebra consists of six operations: addition, subtraction, multiplication, division, maximization, and minimization. The first four operations are defined and denoted as in ordinary algebra. Maximization is defined as the operation of comparing two algebraic quantities and selecting the maximum of the two. Minimization is defined as the operation of comparing two algebraic quantities and selecting the minimum of the two. They are denoted by the symbols $\vee$ and $\wedge$ respectively. $a \vee b$, read a max b, equals a if $a>b$ and b if $a<b$. Similarly $a \wedge b$, read a min b, equals a if $a<b$ and b if $a>b$. By use of these operations on algebraic quantities, the following rules (among others) have been derived

(A). $a \wedge b = -[(-a) \vee (-b)]$

(B). $a+(b \vee c) = (a+b) \vee (a+c)$

The circuit being analyzed and the symbols used are shown in Fig. 8. The positive direction of voltages is as shown. The circuit shown is that for a-c control. However, this analysis will apply equally well for the corresponding d-c controlled circuit for the corresponding signal voltage.

$e_c^{dc} = e_c^{ac}, \; 0 < \omega t < 180°$

$e_c^{dc} = -e_c^{ac}, \; 180° < \omega t < 360°$

Before particularizing to one of the cores, the symbols are used without the core I and II subscripts, i.e., e and e_R refer to the reset-winding and to the control-resistor voltages respectively of either core. In addition, I_m is the maximum magnitude of the reset-circuit magnetizing current; it is assumed constant. $E_R = I_m R_c$, a constant.

Define the magnitude of the maximum volt-second integral, referred to the reset winding, which can be applied to a core during its ghc.

$$L = -\int_{\text{ghc}} \frac{e_{ac}}{N}\,dt = \int_{\text{rhc}} \frac{e_{ac}}{N}\,dt \tag{4}$$

Define the measure of quiescent reset (not applicable to other than the quiescent state)

$$\alpha = \frac{\int_{\text{rhc}} \left(\frac{e_{ac}}{N} - e \right) dt}{L} = \frac{\int_{\text{rhc}} \left(e_R - \frac{k-1}{N} e_{ac} \right) dt}{\int_{\text{rhc}} \frac{e_{ac}}{N}\,dt} \tag{5}$$

where e and e_R are quiescent values $\alpha = 0$ corresponds to full reset, $\alpha = 1$ corresponds to no reset, with a linear relationship between intermediate values of α and intermediate resets.

Consider a control circuit. In the quiescent state during the rhc

$$e_R = e_{ac}' \wedge E_R \tag{6}$$

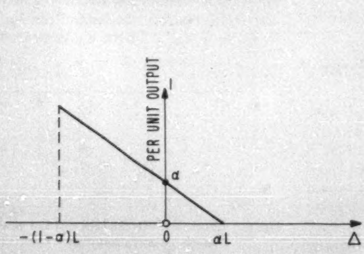
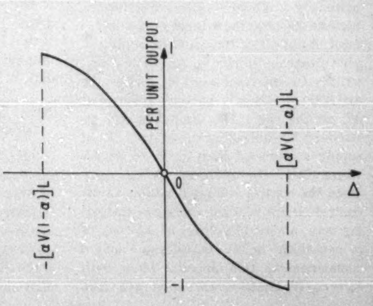


Fig. 9 (A) (left). Half-cycle output of new magnetic servo amplifier versus Δ

Fig. 9 (B) (right). Servo output of new magnetic servo amplifier versus Δ



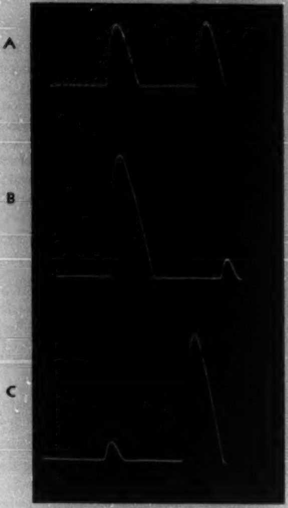


Fig. 10 (A). Steady-state quiescent output current oscillogram. (B) Steady-state output current oscillogram for a positive signal. (C). Steady-state output current oscillogram for a negative signal

$$e=e_{ac}'-e_R$$
$$=e_{ac}'-(e_{ac}'\wedge E_R) \qquad (7)$$

From rule (A)

$$e=e_{ac}'+[(-e_{ac}')\vee(-E_R)]$$

From rule (B)

$$e=0\vee(e_{ac}'-E_R)$$

But the gating circuit limits e to e_{ac}/N the rhc giving

$$e=[0\vee(e_{ac}'-E_R)]\wedge\frac{e_{ac}}{N} \qquad (8)$$

The volt-second integral received by the core during each rhc during the quiescent state is

$$\int_{\text{rhc}} edt=\frac{\int_{\text{rhc}}\frac{e_{ac}}{N}\,dt-\int_{\text{rhc}}\left(\frac{e_{ac}}{N}-e\right)dt}{L}\,L$$

$\int_{\text{rhc}} edt=(1-\alpha)\,L$ in the quiescent state. (9)

Now consider the operation of the amplifier with a control voltage e_c applied. Consider a control circuit during the rhc

$$e_R=(e_{ac}'-e_c)\wedge E_R \qquad (10)$$
$$e=(e_{ac}'-e_c)-[(e_{ac}'-e_c)\wedge E_R]$$

Using rules (A) and (B) just as was done in the quiescent state

$$e=0\vee(e_{ac}'-e_c-E_R)$$

from the control circuit. But the gating circuit still limits e to e_{ac}/N. Therefore

$$e=[0\vee(e_{ac}'-e_c-E_R)]\wedge\frac{e_{ac}}{N} \qquad (11)$$

Now consider the change in the reset-winding voltage caused by the control voltage

$$\Delta e=[0\vee(e_{ac}'-e_c-E_R)]\wedge\frac{e_{ac}}{N}-$$
$$[0\vee(e_{ac}'-E_R)]\wedge\frac{e_{ac}}{N} \qquad (12)$$

For the case where $(e_{ac}'-E_R)\le e_{ac}/N$ and $(e_{ac}'-e_c-E_m)\le e_{ac}/N$ at all times during the rhc, Δe can be simplified to

$$\Delta e=[0\vee(e_{ac}'-e_c-E_R)]-[0\vee(e_{ac}'-E_R)] \qquad (13)$$

The integral of this voltage change over the rhc determines the effect of the control signal on the next half-cycle. Designate this change in the reset voltage-time integral by Δ

$$\Delta=\int_{\text{rhc}}\Delta e dt \qquad (14)$$

$$\Delta=\int_{\text{rhc}}\left\{[0\vee(e_{ac}'-e_c-E_R)]\wedge\frac{e_{ac}}{N}-\right.$$
$$\left.[0\vee(e_{ac}'-E_R)]\wedge\frac{e_{ac}}{N}\right\}dt \qquad (15)$$

For the case where

$$(e_{ac}'-E_R)\le\frac{e_{ac}}{N}\text{ and }(e_{ac}'-e_c-E_R)<\frac{e_{ac}}{N}$$

throughout the rhc

$$\Delta=\int_{\text{rhc}}\{[0\vee(e_{ac}'-e_c-E_R)]-$$
$$[0\vee(e_{ac}'-E_R]\}dt \qquad (16)$$

It is this quantity Δ which when evaluated during one half-cycle determines the output of the amplifier on the following half-cycle.

Consider the quantitative effect of Δ on the following half-cycle. When $\Delta=0$, the amplifier will furnish its quiescent output on the following half-cycle. When $\Delta=-(1-\alpha)L$, the amplifier will furnish maximum output on the following half-cycle. When $\Delta=\alpha L$, the amplifier will furnish zero output on the following half-cycle. A linear relationship, determined by these points, exists between Δ and the amplifier output on the following half-cycle.

In considering the operation of the magnetic servo amplifier, it is the fundamental component of the output which is of interest. Thus, in considering an entire cycle, the quiescent output on both half-cycles will yield no fundamental output. Full output on the first half-cycle and zero output on the next half-cycle will yield maximum fundamental output. Zero output on the first half-cycle followed by maximum output on the next half-cycle will yield maximum reversed fundamental output.

Fig. 11. A transient output current oscillogram

The relationships between Δ and the half-cycle output is shown in Fig. 9(A). If it is assumed that Δ's equal in magnitude but opposite in sign are applied to the two cores on alternate half-cycles, Fig. 9(B) gives the form of the relationship between Δ and the fundamental output.

Description of Laboratory Circuit

Representative circuits of the magnetic servo amplifiers of Figs. 1, 4, 5, and 7 were constructed in the laboratory for a supply frequency of 60 cycles and satisfactorily operated 2-phase motors. Particular attention was paid to the circuit of Fig. 7 with the following components and parameters: Supply frequency of 60 cycles; Deltamax cores no. *5320-D1* with an inside diameter of 1.500 inches, an outside diameter of 2.500 inches, and height of 0.500 inch; $N_g=3{,}000$ turns no. 26 heavy Formvar; $N_c=60$ turns no. 26 heavy Formvar; $N=50$; $e_{ac}=115$ volts rms; $e_{ac}'=190$ volts rms; $R_c=10{,}000$ ohms; *G10* rectifiers; and Ford $2^1/_2$ watt servomotor drawing no. *979646*.

Fig. 10(A) is an oscillogram of the quiescent output current of this amplifier. Fig. 10(B) is an oscillogram of the output current of this amplifier with a 10-volt rms control signal. Fig. 10(C) is an oscillogram of the output current of this amplifier with a 10-volt rms control signal of the opposite phase. Fig. 11 is an oscillogram of the transient response of the amplifier to a suddenly applied 8-volt rms control signal. The upper trace is of the control signal which was suddenly and randomly applied.

Fig. 12(A) gives the 60-cycle component of the output voltage of the amplifier as a function of the control voltage when the amplifier is operating into a 200-ohm resistive load. The voltage gain in the

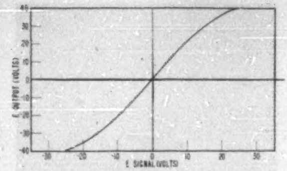


Fig. 12(A). Voltage gain versus input signal for laboratory circuit

range of 0- to 1-volt control signal is about 3; the voltage gain in the over-all range of the amplifier varies from about 3 to about 2. Measurements of the input impedance in the steady state give an input impedance of 20,000 ohms. Thus, the 0- to 1-volt control signal range has a power gain of about 900, and the over-all range of the amplifier from about 900 to about 400. This represents an optimum condition for the laboratory circuit described.

The conditions which bring about this optimum condition should be noted here. Starting with a given circuit with an equal number of control-winding and gating-winding turns, the voltage and power gains are very small, as expected. Decreasing the number of control-winding turns while maintaining the gating-winding turns constant will, at first, increase the voltage and power gains, also as expected. However, this procedure causes a decrease in the control-winding reset voltage and an increase in the control-winding magnetizing current. This, in turn, causes an increase in the control-circuit supply voltage and the various control-resistor and rectifier-voltage drops. Finally, the ratio of control-circuit supply voltage to control-winding voltage becomes so large that the small variations in control resistor and other voltages in the control-circuit become of the same order of magnitude as the control-winding voltage. This, therefore, puts a lower limit on the number of control-winding turns. Below this number, the optimum condition, the voltage and power gains decrease instead of increase.

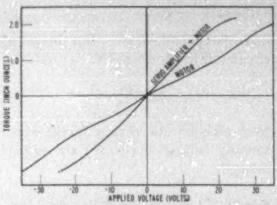


Fig. 12(B). Torque versus applied voltage for servomotor and for servo amplifier and motor

The impedance of the laboratory circuit tested was about 20,000 ohms in the steady state. However, in the dynamic state, the impedance seen by the control source is often that of the control resistor alone, thus necessitating its 10,000-ohm value. In addition, when large signals saturate the core, the control source will see only the control resistor.

Fig. 12(B) gives the locked-rotor torque of the servomotor as a function of the applied voltage. In one case, the applied voltage is applied directly to the control winding of the servomotor; in the other case, the applied voltage serves as the control signal of the amplifier whose output is applied to the control winding of the motor.

Conclusions

The new magnetic servo amplifier for 2-phase servomotor control has the major advantage of half-cycle response. It is simple and gives satisfactory operation for some applications with a single stage. The circuit is operable with either a-c or d-c control, and the control resistors can be varied at will through a wide range, enabling the circuit to be used with a wide variety of control sources. The circuit will or will not furnish a quiescent damping current as desired.

The circuit, however, has several disadvantages. The back electromotive force of the motor when running is detrimental to the operation of the amplifier throughout the entire cycle except when reversing the motor. In addition, the motor must carry the currents of both gating windings. All in all, the advantages outweigh the disadvantages, and it is believed that this circuit will find useful applications in the 2-phase servomotor control field.

References

1. On the Mechanics of Magnetic Amplifier Operation, **R. A. Ramey.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

2. On the Control of Magnetic Amplifiers, **R. A. Ramey.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2124–28.

3. An Improved Magnetic Servo Amplifier, **C. W. Lufcy, A. E. Schmid, P. W. Barnhart.** *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 281–89.

4. A New Type of Magnetic Servo Amplifier, **W. A. Geyger.** *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 272–80.

5. Magnetic Amplifier Circuits, **W. A. Geyger.** McGraw-Hill Book Company, Inc., New York, N. Y., p. 181, Fig. 12.6.

6. A Rectifier Algebra, **D. H. Schaefer.** *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 679–82.)

A High-Accuracy Static Time Delay Device Utilizing Transistors

G. F. PITTMAN, JR.
ASSOCIATE MEMBER AIEE

THE fixed voltage-time integral value determined by the swing of the flux level of a rectangular hysteresis loop magnetic core from one saturation level to the other has been exploited in a variety of ways, e.g., in magnetic amplifiers, saturating transformers, etc. Nevertheless, extensive utilization of the integrating capabilities of cores has been limited by the functions which associated circuit components are capable of performing. Employing junction transistors as circuit components, however, opens new possibilities of using this characteristic of saturable cores to perform a variety of useful functions.

This paper describes circuitry employing junction transistors as nearly ideal, controlled switches in conjunction with rectangular hysteresis loop magnetic cores to perform pulse shaping and preset pulse-counting functions. By deriving pulses from a fixed frequency source, the circuitry may be designed to yield quite long time delays with extreme accuracy, because of the digital nature of operation employed. The circuitry and the manner in which transistors are used permits operation under conditions of relatively wide supply voltage variations and am-

---

Paper 55-74, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing December 3, 1954.

G. F. Pittman is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

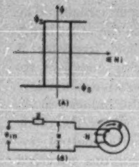


Fig. 1. A—Idealized rectangular hysteresis loop B—Simple saturable core circuit

bient temperatures extending to at least 80 degrees centigrade (C) without adversely affecting the timing accuracy.

Principles of Operation

Tape-wound toroidal cores of 50-50 nickel-iron alloy exhibit very nearly the idealized rectangular hysteresis loop shown in Fig. 1(A). Essential features of this magnetic characteristic are the sharp saturation at some definite core flux, ϕ_s, and the retention of the existing core flux level upon removal of all applied magnetomotive forces.

Let us suppose that such a core is provided with a winding and a voltage applied as in Fig. 1(B). The voltage $e = N(d\phi/dt)$ is the voltage "absorbed" by the core in changing its flux level in response to the applied voltage e_{in}. Change in core flux level can therefore be related to the volt-time integral of the voltage e

$$\Delta\phi = \frac{1}{N}\int_{\Delta t} edt$$

From this it is seen that, if core flux is changed from negative to positive saturation, a volt-time area must be supplied to the core equal to

$$\int edt = 2N\phi_s$$

If the voltage drop across the series impedance element Z of Fig. 1(B), which includes winding resistance as well as any other series impedances, is negligible compared to e_{in} and e, then

$$e_{in} = e$$

and

$$\int e_{in}dt = 2N\phi_s$$

Thus, if a low-resistance winding on the core is supplied with voltage from a low-impedance source, the core (assuming that it is driven from negative to positive saturation) integrates the input voltage with respect to time until a definite value of the integral is attained, at which time it saturates sharply. Because of the retentivity properties of rectangular loop cores, the input voltage need not be continuous but may be applied in the form of pulses; in this case the core retains the flux level at the end of one pulse until the next pulse is applied, and the integration process takes place as it would with a continuous voltage applied. Thus, for pulse input voltages

$$\int_{\Delta t_1} edt + \int_{\Delta t_2} edt + \int_{\Delta t_3} edt + \ldots + \int_{\Delta t_n} edt = 2N\phi_s$$

where $\int_{\Delta t_n} edt$ denotes the volt-time integral value of the nth pulse. If the input pulses are of fixed volt-time integral value; i.e., if

$$\int_{\Delta t_1} edt = \int_{\Delta t_2} edt = \int_{\Delta t_3} edt = \ldots = \int_{\Delta t_n} edt$$

then

$$n\int_{\Delta t_n} edt = 2N\phi_s$$

or

$$n = \frac{2N\phi_s}{\int_{\Delta t_n} edt}$$

so that a definite number of input pulses are required to saturate the core. It is this voltage integrating property of the rectangular loop saturable core, particularly with pulse inputs, which forms the basis of operation of the device to be described.

If a core is initially placed in the negative saturation state, and the saturation of the core in the positive direction is detected and an output signal immediately initiated, a variety of useful functions can be performed by virtue of the integrating property of the core. For instance, if the input voltage of Fig. 1(B) varies in a random manner with time, an output signal will be obtained when the time integral of the input voltage attains some fixed value as determined by the saturation flux of the core and the number of turns of the input winding. On the other hand, if e_{in} consists of a series of unidirectional pulses of equal volt-time area; then, as stated before, a fixed number of such pulses will be required to drive the core from negative to positive saturation, and an output signal will be obtained after application of a preset number of pulses. If such input pulses are supplied synchronously with a fixed frequency source; e.g., an a-c power source, a given number of pulses will represent a fixed time interval, and an output signal will be delivered after a preset time interval.

Since, in general, it is desirable to reset the core to negative saturation immediately after it has been driven to positive saturation, a convenient means of obtaining an output signal is provided, since the voltage induced in an output winding on the core during resetting serves ideally as an output pulse. Such a pulse has a fixed volt-time area determined by the core saturation flux and the number of turns of the output winding; thus, it is ideally suited to drive another core, operating in a similar manner, partially toward positive saturation. By such cascading, the range of integrating or counting can be greatly extended.

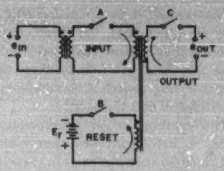


Fig. 2. Basic preset counter circuitry

Although the principle of operation of a magnetic core integrator, as just set forth, is quite elementary and straightforward, the circuitry required to reduce the principle to practice is quite specialized in its required functions.

Basic Circuitry

Basically three circuits are required. These are:

1. An input circuit and winding providing a means of applying input voltage pulses to the core.
2. A reset circuit and winding providing a means for resetting the core to negative saturation upon saturation in the positive direction.
3. An output circuit and winding providing a means of applying the induced voltage during resetting to a load or succeeding stage.

These circuits are shown schematically in Fig. 2, each incorporating a switch which it is assumed may be opened and closed at appropriate times.

During an input voltage pulse, it is desired that the entire input voltage be translated into core flux change. This dictates that the input circuit possesses low series impedance other than the inductive impedance of the core and that current flow in the input circuit be kept to a minimum. The latter condition implies that there be no loading on the core due to other circuits' linking it. Therefore, during input pulsing, it is required that switch A be closed and switches B and C be open.

During resetting it is desired that the resetting battery voltage be applied to the resetting winding and that the voltage induced in the output winding be transmitted to the output terminals. Since the input circuit has been specified to be of very low impedance, it is necessary that switch A be opened during resetting in order that the input circuit shall not act as a short-circuited winding linking the core. It follows then that, during resetting, switch A should be open and switches B and C closed.

During intervals of zero input voltage, switches A and C may be either open or closed since no voltages exist in these circuits during such intervals. Switch B, however, must be open since the resetting battery would otherwise be applied to the resetting winding causing premature resetting to take place.

It should be noted that switch C is required to permit all current flow in one direction and block all current flow in the opposite direction; therefore, a simple diode will serve adequately the purpose of this switch. Switches A and B, however, are required both to block and to permit current flow in the same direction, depending upon the circuit operation being performed at a given time. Simple diode elements will therefore not perform the functions required of switches A and B.

A Practical Pulse-Integrating Circuit

In the circuit shown in Fig. 3, transistors A and B perform the functions of switches A and B respectively of Fig. 2, while the diode in the output circuit performs the function of switch C. In this circuit, input voltage pulses are supplied through a transformer having two secondary windings in order to both apply voltage to the input circuit and also to apply, simultaneously, the appropriate driving voltage to the base of transistor A.

It has been shown[1] that a junction transistor can function as a nearly ideal switch between its emitter and collector electrodes if appropriate driving voltages and currents are applied to the base electrode. The pertinent conditions may be summarized as follows (for *P-N-P* transistors).

1. With the base positive with respect to both the emitter and the collector, the transistor acts as an open switch, permitting only a very small leakage current to flow in the emitter-collector circuit.
2. With the base negative with respect to either the emitter or the collector, the transistor acts as a closed switch, offering very low resistance to current flow in the emitter-collector circuit.

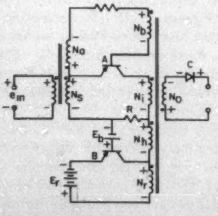


Fig. 3. Preset counter circuit during an input pulse

During application of an input pulse the polarities of voltages existing on the various windings of the transformer and saturable core are as shown in Fig. 3. The voltage induced in winding N_a is of such polarity as to drive the base of transistor A negative with respect to its emitter and thus permit current flow from emitter to collector. The input pulse voltage induced in N_s is thus applied directly to winding N_i of the saturable core, the value of resistor R being such that the voltage drop of the core magnetizing current is negligible. As a result of the voltage applied to N_i, voltages are induced in windings N_b and N_h as shown. The voltage induced in N_b adds to the existing voltage applied to the base of transistor A to render it conducting, while the voltage induced in N_h adds to the bias voltage E_b applied to the base of transistor B to hold it blocking. The output diode, meanwhile, blocks the voltage induced in N_0 from the output terminals. Thus the conditions specified for the input pulsing interval are fulfilled, i.e., switch A is closed while switches B and C are open.

During intervals between input pulses, no voltages exist on any of the windings of the saturable core. During this time the bias battery in the base circuit of transistor B holds it blocking, thus preventing the application of the resetting voltage E_r to winding N_r. It may be noted that any negative input pulses will be blocked from the input winding by transistor A since the voltage induced in N_a will be of a polarity to block A rather than rendering it conducting as is the case with a positive input pulse.

Upon saturation of the core during some input pulse, the induced voltages in the saturable core windings disappear and the pulse voltage appears across resistor R. This voltage, which also appears in the base circuit of transistor B, overcomes the bias battery voltage, placing a negative potential on the base of B and causing it to cease blocking. The resetting voltage E_r is thus applied to N_r and resetting of the core is initiated. During resetting the polarities of voltages existing in the circuit are as shown in Fig. 4.

The voltage induced in N_b is of such a polarity as to place a positive potential on the base of transistor A, thus blocking current flow in the input circuit. The voltage induced in N_h, on the other hand, is of the polarity to hold transistor B conducting, thus insuring complete resetting of the core to negative saturation. The output voltage pulse induced in N_0 during resetting is passed by the diode and a signal appears at the output terminals. Thus, all conditions specified for the resetting interval are met, i.e., switch A is open while switches B and C are closed.

It may be noted in Figs. 3 and 4 that the resetting transistor is used in an inverted connection, i.e., when acting as an open switch the collector is positive with respect to the emitter. This connection is used here in preference to the normal one to minimize leakage current during blocking which might otherwise cause partial resetting in the intervals between input pulses, and result in inaccurate counting.

A Fixed Time Delay Device

As pointed out previously, the pulse-integrating circuit functions as a preset counter if supplied with input pulses of constant volt-time area, and a preset time delay if these pulses are synchronous with a fixed frequency source. A saturating transformer driven from a fixed frequency source would provide such pulses; however, to provide more nearly ideal switching of the transistors, a rectangular input pulse is desirable.

The pulse-shaping circuit shown in Fig. 5 produces, in response to an input pulse, a rectangular output pulse of fixed volt-time area. Suppose that initially the saturable core has been reset to negative saturation under the influence of E_p acting on winding N_r through resistor R. The transistor is held in the blocking state by the bias voltage E_b applied to its base. An input pulse of the polarity shown overcomes the bias voltage rendering the transistor conductive and applying E_p directly to winding N_p. Since the transistor emitter-collector circuit represents a much lower impedance than R

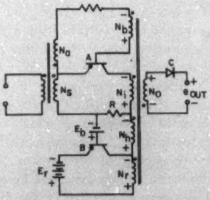


Fig. 4. Preset counter circuit during resetting

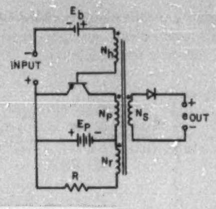


Fig. 5 (left). Pulse shaper circuit

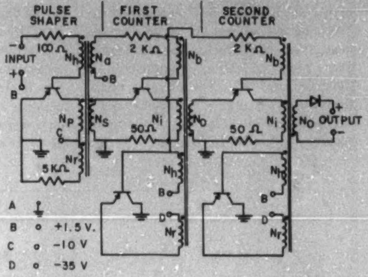


Fig. 6 (right). Pulse shaper and 2-stage preset counter circuit

(on a common turns base), the core is driven toward positive saturation, the voltage induced in winding N_h acting to hold the transistor conducting until the core saturates. The rectangular voltage pulse induced in the output winding N_s during this time has a fixed volt-time area determined by the saturation flux of the core and the number of output winding turns. Upon saturation of the core the induced voltages disappear and again the transistor blocks. The core is then reset through the N_r circuit in preparation for the next input pulse, the diode in the output circuit blocking the induced voltage during resetting from the output terminals. Thus, in response to a rather arbitrary input pulse, the pulse-shaping circuit delivers a unidirectional rectangular pulse of fixed volt-time area.

In combining a pulse-shaping circuit and a number of cascaded counter stages, output windings of the preceding stage serve as the transformer secondary windings shown as the input in Figs. 4 and 5. Furthermore, the output diode is necessary only in the last stage since the input transistor of the succeeding stage serves this purpose as well as its other switching functions. In such a series of cascaded staegs, each is isolated, by virtue of the solely magnetic coupling, from the others. Therefore, regardless of the number of stages, only one common d-c power supply is required.

The use of a saturable-core pulse shaper and the means employed to derive output pulses from the counter stages provide an important stability feature when a pulse shaper and several stages are operated in cascade, e.g., in order to obtain long time delays. The saturation flux of saturable magnetic cores varies somewhat with temperature; this variation becomes quite appreciable when wide ambient temperature variations, e.g., −65 to +70 C are considered. Thus, if input pulses of truly constant volt-time area are applied to a counter stage, the number of pulses required to drive the core from negative to positive saturation will vary with ambient temperature, introducing inaccuracy in the time delay. However, by using the swing of a core from saturation to saturation to determine the volt-time area input to each stage, as is done in the circuitry described, no inaccuracy will be introduced if all cores (pulse shaper and counter stages) vary in a like manner with temperature. Since, in general, all cores will be of the same material, this will be quite nearly the case. Furthermore, since transistors are used in such a manner that they exhibit quite negligible voltage drops when conducting and very small leakage currents when blocking, wide variations in transistor characteristics, both between units and with temperature, may be tolerated without affecting the operation of the circuit. The net result of these considerations is that accurate and reliable operation, unaffected by wide ambient temperature variation, is obtained.

Experimentation

Time delay units have been built up consisting of a pulse shaper and two cascaded counter stages as shown for example in Fig. 6. In this unit a somewhat different method of detecting saturation of the counter core is employed than shown in Figs. 3 and 4. Instead of using the voltage developed across a series resistor upon saturation of the core to trigger the resetting transistor, the voltage induced in the resetting transistor holding winding (N_h of Figs. 3 and 4) as the core flux drops from its maximum saturated value back to the residual value at the end of the input pulse is employed. Furthermore, the function of winding N_a of Figs. 3 and 4 is performed in the circuit of Fig. 6 by other windings in addition to their other functions. On the pulse-shaper core the series combination of windings N_s and N_a is used, while on the first counter core windings N_0 and N_h are used. This is possible since the method of saturation detection employed here eliminates the necessity for the resistor common to the input circuit and the resetting transistor base circuit of a counter stage, also enabling the common bias voltage to be used as a blocking bias in the base circuit of the counter stage input transistor of Fig. 6. This bias voltage ensures more complete blocking of the input transistor, especially under conditions of high ambient temperature. Despite these changes, however, the operation of the circuit is essentially the same as that described in the preceding section.

The unit shown in Fig. 6 is designed to require 10 input pulses to saturate each stage; thus 100 input pulses are required to saturate the last stage and yield an output pulse. The saturable cores used in the circuit are all wound toroids of 2-mil Hipernik V material. All transistors are Westinghouse *P-N-P* junction type *2N56*.

Using input pulses derived from the 60-cycle power line, the unit was checked for operation over an ambient temperature range of −65 to +80 C. In all cases 100 input pulses were required to obtain an output pulse from the second stage. Because of the digital nature of the device, it is expected that such absolute accuracy would be carried through to units of many more stages and, consequently, much longer time delays.

To check the retention of information (i.e., existing counter core flux level) over long periods of time, input pulses were turned off after less than 100 were supplied, and then turned on again after intervals up to 8 hours. It was found that, despite the elapsed time, the proper total number of input pulses produced an output signal. These results indicate that, by cascading stages, time delays in excess of 8 hours should be obtainable with an accuracy of considerably less than 1 half cycle of the input pulse frequency.

Conclusion

Circuitry employing junction transistors and rectangular hysteresis loop magnetic cores has been described which enables effective utilization of the voltage integrating capabilities of the core to perform preset counting operations. An example has been given of an application of such circuitry in a fixed time delay device to provide very long delays to an accuracy of a small fraction of the period of the fixed frequency source from which it is driven.

The components and circuitry which are employed are such that reliable operation is obtained over wide ambient temperature ranges and with wide tolerances on power supply voltage variations. The device is expected to display the excellent reliability and life which are characteristic of its components, junction transistors and saturable magnetic cores.

Reference

1. Junction Transistors Used as Switches, R. Louis Bright. *AIEE Transactions*, vol. 74, pt. I, Mar. 1955, pp. 111–21.

A New Audio Telegraph Carrier Terminal

C. A. HIGGINS
NONMEMBER AIEE

E. A. GILBERT
ASSOCIATE MEMBER AIEE

A NEW telegraph carrier terminal has been developed to meet the ever-increasing demand for telegraph communication facilities. The terminal is designed to meet the frequency spectrum requirements for 12-channel telegraph systems.[1] These standards were set up to allow the telegraph communication systems of the U. S. Army, Navy, and Air Force and many existing commercial systems to be integrated for joint use in the interest of national defense.

Application

Normal center frequencies of the telegraph channels are at 425, 595, 765, 935, 1,105, 1,275, 1,445, 1,615, 1,955, 2,380, 2,805, and 3,230. The first eight channels are 170 cycles wide with the space frequency 42.5 cycles per second (cps) below center-band frequency and the mark frequency 42.5 cps above center-band frequency. The last four channels are 425 cps wide with the space frequency 85 cps below center-band frequency and the mark signal 85 cps above center-band frequency. The tolerance allowed on the mark and space frequencies as generated is ±3 cps. The narrow-band channels are capable of transmitting telegraph messages up to 100 words per minute. Broad-band channels transmit high-speed messages up to a 250-words-per-minute rate or carry four time-division multiplex telegraph channels at 60 words per minute.

The equipment described in this paper is used for fixed-plant installations under conditions of continuous operation over extended periods of time. Reliability, long life, simplicity of design, and ease of maintenance are the guiding considerations for equipment of this type. Front panel test jacks are provided for system trouble-shooting and initial installation adjustments. Both land lines and radio links are used for connecting the transmitting and receiving sections. For long-distance radio links, the equipment may be switched into a frequency diversity system and, with slight modification, can be used for space diversity operation. Fig. 1 shows the complete 12-channel terminal with the transmitting units mounted in the cabinet at the right and the receiving units mounted in the cabinet at the left. The individual chassis are $5\frac{1}{4}$ inches high and may be withdrawn from the cabinet and turned up for maintenance, as shown in Fig. 2.

Because this equipment has a wide field of application, many options are required in electric connections and circuitry. Plug-in frequency-determining components are used to permit the exchange of basic chassis units for maintenance purposes and to permit future conversion of the equipment to a 16-channel system or other systems, as may be desired. Fig. 3 shows a complete filter set for one telegraph channel.

Group modulator or frequency converter units are provided for converting the frequencies of the narrow-band channels into the frequency spectrum of 1,742.5 to 3,017.5 cps for transmission with the normal narrow-band channels to make a 16-channel system. The 1,742.5- to 3,017.5-frequency spectrum is modulated again at the receiving end and converted to the normal channel frequencies.

Systems

A block diagram of a typical terminal system with 12 telegraph channels is shown in Fig. 4. Twelve sending teletype printers and twelve receiving teletype printers are required. The communication link may be either a land line or a radio circuit. In the output of each transmitter unit is a 20-decibel (db) band-pass filter that can be operated in parallel with others of the same type. These filters attenuate frequencies 20 db or more in the adjacent channels. Outside the passband the shunt reactance is high. Only the amplifiers of the transmitter and receiver frequency converters are used in this connection. The input to each receiver unit is a 40-db band-pass filter that is operated in parallel with the other 11. An application of the frequency converters in a 16-channel system is shown in Fig. 5. The channels 9 through 16 are modulated in the transmitter frequency converter against a precise 3,400-cps crystal-controlled standard frequency and converted to the 1,742.5-cps-to-3,017.5-cps frequency range. After transmission, these frequencies are again modulated against a 3,400-cps standard frequency and are returned to their normal frequencies.

Transmitter Unit

The transmitter group contains 12 telegraph transmitter units. These units convert the d-c telegraph pulses in the loop circuit to frequency shift "mark" and "space" audio signals for each channel. Each transmitter consists of a 15-kc loop oscillator, a d-c amplifier with tele-

---

Paper 55-94, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 25, 1954; made available for printing November 18, 1954.

C. A. Higgins and E. A. Gilbert are with the Radio Frequency Laboratories, Inc., Boonton, N. J.

Development of this terminal was carried on for the military under a Navy Department, Bureau of Ships contract and is the result of co-operative efforts of many persons in the Bureau of Ships, Signal Corps of the U. S. Army, and the Radio Frequency Laboratories, Inc. The military nomenclature for the equipment is telegraph carrier terminal *AN/FCC-3*.

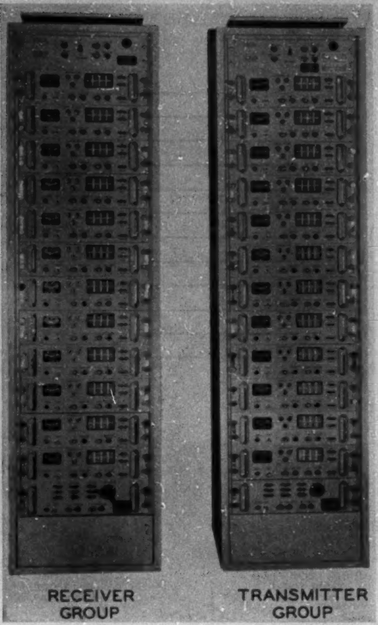


Fig. 1 (left). Twelve-channel audio carrier telegraph terminal

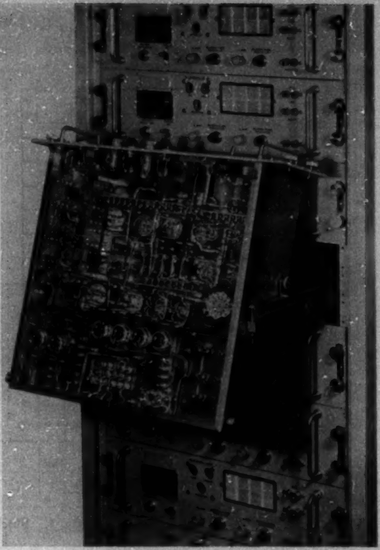


Fig. 2 (right). Receiver unit in maintenance position

graph bias controls, a plug-in frequency-shift tone oscillator, an output amplifier, a plug-in 20-db band-pass filter, and power supplies. The arrangement of these circuits is shown in the block diagram of Fig. 6. The 15-kc loop oscillator isolates the d-c loop circuits from the transmitter chassis which is at ground potential. The 15-kc signal is rectified and applied to a d-c amplifier. Telegraph bias correction controls are designed into the d-c amplifier. The tone oscillator is a resistance stabilized Hartley circuit, the frequency of which is shifted electronically by adding capacitance to the tuned circuit with a diode switch. Two power supplies are included in the unit; a 120-volt d-c supply for the telegraph loop circuit, and a 216-volt d-c supply for the amplifiers and oscillators. The loop-circuit current can be supplied from either the internal power supply or external battery, and is selected by a switch in the equipment. The unit operates from a single-phase power source of 115/230 volts, 50/60 cycles. Controls are included for loop-current adjustment and for signal output level. Mark and space output levels are within 1 db and can be adjusted over a range from −35 to +6 db referred to 1 milliwatt.

Receiver Unit

As shown in Fig. 4, the receiver group contains 12 telegraph carrier receiver units. These units convert the frequency shift audio telegraph signals of each channel to d-c pulses for operation of a telegraph printer or other telegraph and equipment. As shown in Fig. 6, each telegraph carrier unit consists of a plug-in 40-db band-pass filter, a limiting amplifier, a plug-in discriminator, a d-c amplifier, and a polar or neutral output circuit. A conventional power supply is included to operate the vacuum-tube circuits as well as to supply the loop current. By throwing a switch, an external battery may be used for the loop-circuit current. Because of the signal-fading conditions normally encountered on radio circuits, the limiting amplifier is designed to operate between +6 and −40 db referred to 1 milliwatt.

Frequency Diversity

Frequency diversity connections are made by switches on each chassis. Fig. 7 shows a block diagram of two receiver units in frequency diversity. The channel center-band frequencies of diversity pairs are spaced apart by 500 or more cycles for two reasons: a selective signal fade on a radio circuit will at times attenuate one of the carrier channels 40 db or more and not affect another signal spaced 500 cps away at the same instant; and the passbands of the two discriminators must not overlap or telegraph bias

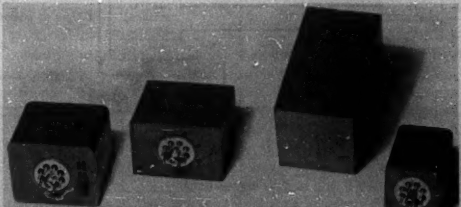


Fig. 3. Plug-in filter set for one channel

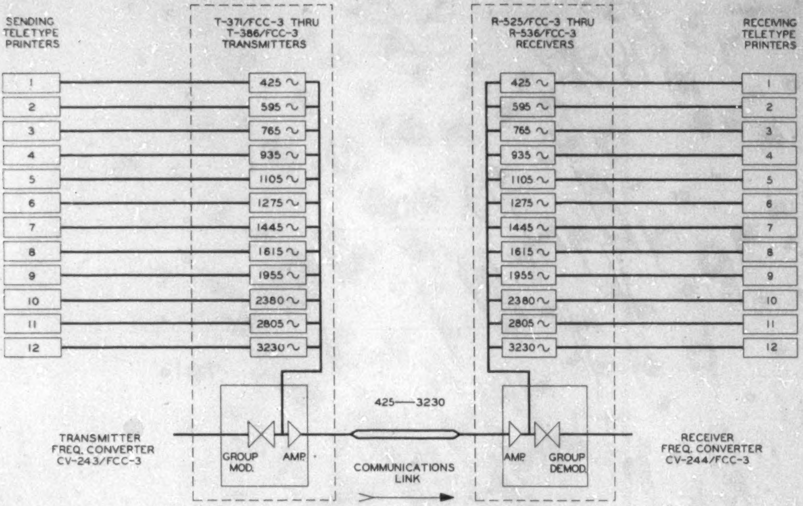

SENDING TELETYPE PRINTERS
T-371/FCC-3 THRU T-386/FCC-3 TRANSMITTERS
R-525/FCC-3 THRU R-536/FCC-3 RECEIVERS
RECEIVING TELETYPE PRINTERS
1
2
3
4
5
6
7
8
9
10
11
12
425 ~
595 ~
765 ~
935 ~
1105 ~
1275 ~
1445 ~
1615 ~
1955 ~
2380 ~
2805 ~
3230 ~
425—3230
TRANSMITTER FREQ. CONVERTER CV-243/FCC-3
GROUP MOD.
AMP.
COMMUNICATIONS LINK
AMP.
GROUP DEMOD.
RECEIVER FREQ. CONVERTER CV-244/FCC-3

Fig. 4. Block diagram of 12-channel connection

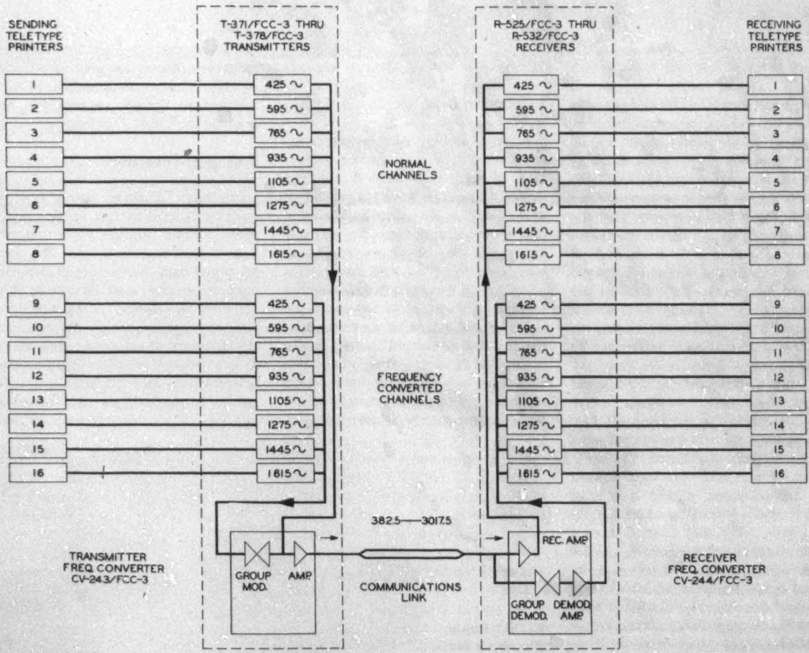

SENDING TELETYPE PRINTERS
T-371/FCC-3 THRU T-378/FCC-3 TRANSMITTERS
R-525/FCC-3 THRU R-532/FCC-3 RECEIVERS
RECEIVING TELETYPE PRINTERS
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
425 ~
595 ~
765 ~
935 ~
1105 ~
1275 ~
1445 ~
1615 ~
NORMAL CHANNELS
FREQUENCY CONVERTED CHANNELS
382.5—3017.5
TRANSMITTER FREQ. CONVERTER CV-243/FCC-3
GROUP MOD.
AMP.
COMMUNICATIONS LINK
REC. AMP.
GROUP DEMOD.
DEMOD AMP.
RECEIVER FREQ. CONVERTER CV-244/FCC-3

Fig. 5. Block diagram of 16-channel connection

Vol. 7
Pt. 1
1955

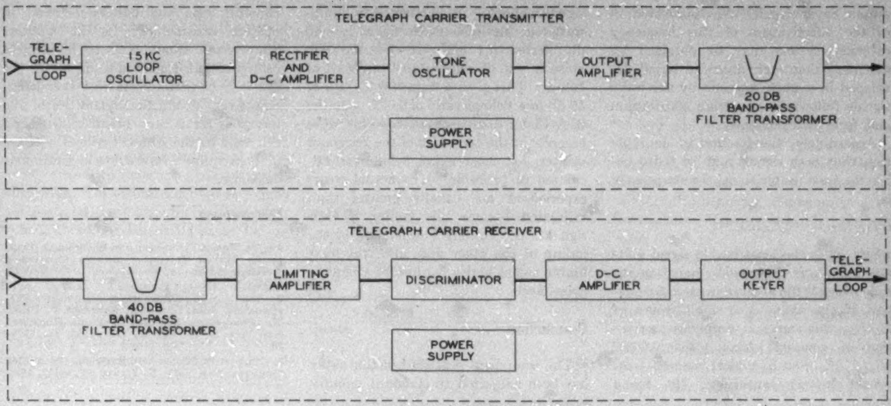

TELEGRAPH CARRIER TRANSMITTER
TELE-GRAPH LOOP
15 KC LOOP OSCILLATOR
RECTIFIER AND D-C AMPLIFIER
TONE OSCILLATOR
OUTPUT AMPLIFIER
20 DB BAND-PASS FILTER TRANSFORMER
POWER SUPPLY
TELEGRAPH CARRIER RECEIVER
40 DB BAND-PASS FILTER TRANSFORMER
LIMITING AMPLIFIER
DISCRIMINATOR
D-C AMPLIFIER
OUTPUT KEYER
TELE-GRAPH LOOP
POWER SUPPLY

Fig. 6. Block diagram of one carrier telegraph channel

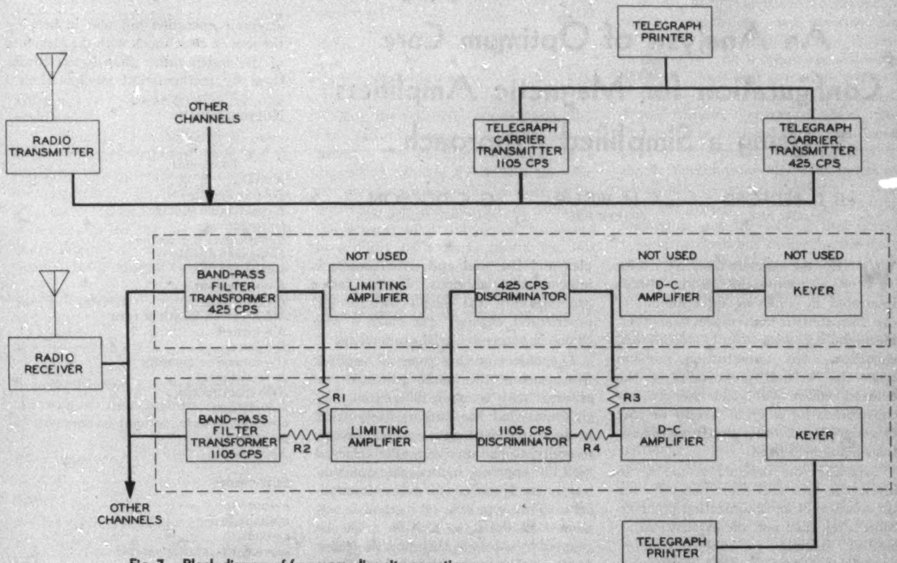

TELEGRAPH PRINTER
RADIO TRANSMITTER
OTHER CHANNELS
TELEGRAPH CARRIER TRANSMITTER 1105 CPS
TELEGRAPH CARRIER TRANSMITTER 425 CPS
NOT USED
NOT USED
NOT USED
BAND-PASS FILTER TRANSFORMER 425 CPS
LIMITING AMPLIFIER
425 CPS DISCRIMINATOR
D-C AMPLIFIER
KEYER
RADIO RECEIVER
R1
R3
BAND-PASS FILTER TRANSFORMER 1105 CPS
R2
LIMITING AMPLIFIER
1105 CPS DISCRIMINATOR
R4
D-C AMPLIFIER
KEYER
OTHER CHANNELS
TELEGRAPH PRINTER

Fig. 7. Block diagram of frequency diversity operation

distortion is introduced when one signal fades out. In the frequency diversity system used, one telegraph printer keys two telegraph carrier transmitter units. The block diagram of Fig. 7 shows the arrangement of two typical units with center-band frequencies at 425 and 1,105 cps. The audio output of both telegraph carrier transmitters is transmitted over a radio link to the receiving site. Audio signals from the output of the radio receiver are selected from the common bus with band-pass filters. The signals are added with equal value resistors *R1* and *R2*, and applied to the first grid of a common limiting amplifier. When one signal frequency is twice the amplitude of the other, the output of the limiting amplifier contains the frequencies of the greater channel. The discriminator of the channel with the greater signal level then controls the d-c amplifier and output keyer circuit. When both signal levels at the input amplifier are equal, the telegraph signal level at the input to the d-c amplifier is about ±5 per cent greater than when the signals are unequal. Limiting action of the d-c amplifier prevents this change of telegraph signal level from creating telegraph distortion in the output keyer circuit. An approximation of the effectiveness of this frequency diversity system may be obtained by assuming that each diversity signal developed in a single antenna by sky wave signals follows the Rayleigh distribution and fade independently.[2,3]

Accordingly, the fraction of unusable time that both signals may be faded below the least usable signal simultaneously is

$$t/T_0=[1-\epsilon^{-0.693(E_L/E_M)^2}]^2$$

where E_L is the least usable signal voltage and E_M is the median signal amplitude which is the amplitude exceeded one-half of the time. For small values of E_L/E_M, this equation simplifies to fraction of unusable time $t/T_0 \approx (0.693)^2 (E_L/E_M)^4$. For a typical system with equal channel sensitivity, the mean signal is about −10 db referred to 1 milliwatt and the least usable signal is −40 db referred to 1 milliwatt when the peak noise is −40 db referred to 1 milliwatt or below. This gives a value for E_L/E_M of 30 db or a voltage ratio of 0.032. On the basis that a 10-millisecond fade may cause an error in the operation of the telegraph printer, one error every 5 hours is calculated as probable. The actual errors experienced are usually greater than estimated because the fading of each signal is not entirely independent of the fading of the other and, also, the least usable signal varies because of changing noise levels.

Conclusion

The equipment described in this paper has been subjected to standard military climatic tests and has been found to operate satisfactorily in the ambient temperature range of 0 to 50 degrees centigrade at humidities up to 95 per cent. Performance has met the design objective. Production quantities of the terminals have been manufactured and delivered to the Armed Services. Many of these equipments are now in continuous daily use.

References

1. Technical Standards for 12-Channel Telegraph System. National Military Research & Development Board, Washington, D. C., June 15, 1951.

2. Evaluation of Polarization Diversity Performance, John L. Glaser, Lawrence P. Faber, Jr. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 41, Dec. 1953, pp. 1774–78.

3. Theory of Sound, Lord Rayleigh. Macmillan Company, New York, N. Y., 2nd ed., vol. I, 1894, p. 41.

An Analysis of Optimum Core Configuration for Magnetic Amplifiers Using a Simplified Approach

B. D. BEDFORD
FELLOW AIEE

C. H. WILLIS
FELLOW AIEE

G. C. DODSON
ASSOCIATE MEMBER AIEE

When the engineer designs a device or a component he is always interested in achieving an optimum of one characteristic with respect to another. He might, e.g., wish to build a cylindrical capacitor. His optimization problem might be to find proportions of the cylinder which will yield the greatest capacitance for a given weight or cost of materials, all other design considerations being held fixed.

Typically, the designer is unable to invest the time and effort necessary for the solution of an optimization problem. More often than not, the explicit mathematical solution for even common, practical problems is prohibitively complex and the trial and error process is interminably laborious. In the latter process, only the highly skilled and experienced engineer can make a wise choice in successive trial proportions.

The object of this paper is twofold: first, to show the use of a little-known criterion with common differential principles in what the authors have found to be a simplified optimization process; secondly, to present the results achieved with this approach in analyzing optimum reactor configurations for magnetic amplifiers. The approach or method is explained in detail so that it might be adapted by others to their own problems. A practical example, one of the difficult cases in the magnetic amplifier analysis, is detailed in Appendix IV. A mathematical proof of the criterion is presented in Appendix I.

It is the little-known criterion, based on linear dimension relationships, that is the crux of simplifying the optimization process. The process or method is referred to here as the "Application of Infinitesimal Per-Unit Increments." Its simplicity comes not only from reduced algebraic operation but also in keeping the user in close touch with the elements of the design rather than isolated from them by mathematical manipulations.

Nomenclature

r, t, a, b, h, w, x, y, z, l_1, l_2, and l_3, etc. = dimensions
A = area
B = flux density
E = generated voltage
f = frequency
G_s = d-c gain per second
i = current density
I = current
$K_{(\,)}$ = constant
L = inductance
MLT = mean length of turn
N = turns
P = power
Q = volume or quantity
R = resistance
SF = space factor
V = voltage
$\mathbf{u}$ = finite, minute, per-unit increments
ρ = resistivity
Δ = increment

Subscripts

o = output
a = apparent
c = control
ca = copper area (apparent)
fe = iron
i = input
L = load
s = supply or surface
t = total or enclosure
$\underline{w}$, $\underline{x}$, $\underline{y}$, $\underline{H}$, and $\underline{D}$: bar under quantity indicates it is a variable

The Method—Application of Infinitesimal Per-Unit Increments

It is quite common in engineering designs to make a change in an existing,

---

Paper 55-68, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31-February 4, 1955. Manuscript submitted March 23, 1954; made available for printing November 26, 1954.

B. D. Bedford and G. C. Dodson are with the General Electric Company, Schenectady, N. Y., and C. H. Willis is with Princeton University, Princeton, N. J.

The authors gratefully acknowledge the contributions made to this study by Dr. H. Poritsky, R. W. Kuenning, H. L. Kellogg, and E. B. Johansson.

Vol.
Pt.
195

designed component in an effort to improve it in some way. In the use of infinitesimal per-unit increments, the same general procedure applies but with greater resolution of purpose and less practical expense. There are five major steps involved:

1. In the first step it is necessary to identify the two characteristics which will be used to determine the optimum, i.e., maximum volume for a given weight of material, best gain for a given cost, etc. These characteristics must be expressed in terms of the physical dimensions and material design parameters of the device. The expressions obtained will be homogeneous polynomials, i.e., each term of the polynomial is the product of the same number of dimensions—homogeneous in degree. For a cylinder where the maximum volume for minimum surface is desired, volume V and surface S are expressed as

$$V=\frac{\pi D^2H}{4} \qquad (1)$$

$$S=2\left(\frac{\pi D^2}{4}\right)+\pi DH$$

These are the characteristics for optimization.

2. Secondly, it is necessary to determine the power of linear dimension to which the two optimizing characteristics are responsive. In the case of surface area, as mentioned in the foregoing, it is responsive to the second power of linear dimension—each term is proportional to the product of two dimensions or

$$S=\frac{\pi D^2}{2}+\pi DH \propto l^2 \qquad (2)$$

$$V=\frac{\pi D^2H}{4} \propto D^2H \propto l^3 \qquad (3)$$

In general terms, the design characteristics will be homogeneous polynomials of the type

$$P=\sum_{i=1}^{n} A_i w^{p_i} x^{q_i} y^{r_i} \qquad (4)$$

where

$$p_i+q_i+r_i=K \text{ (degree of each term)} \qquad (5)$$

3. In the third step, a shape for the device is assumed. For this shape, infinitesimal per-unit increments are applied to individual dimensions and the per-unit incremental changes in the optimizing characteristics are noted. This is done for each dimension in turn. This step is a simple algebraic operation the simplicity of which comes from the principles of differentials. Take any term T of the polynomial P such as

$$T=A_1 w^{p_1} x^{q_1} y^{r_1} \qquad (6)$$

Apply increments to each variable w, x, and y and note the effect on the dependent variable T

$$\Delta T_{\underline{w}}=\frac{\partial T}{\partial w}dw=\frac{\partial T}{\partial w}w\frac{dw}{w}$$

$$=A_1p_1w^{(p_1-1)}x^{q_1}y^{r_1}w\frac{dw}{w}=p_1T\frac{dw}{w} \qquad (7)$$

The actual value of $(dw)/w$ being quantity $a\mathfrak{u}$

$$\Delta T_{\underline{w}}=p_1Ta\mathfrak{u} \qquad (8)$$

Similarly

$$\Delta T_{\underline{x}}=q_1Tb\mathfrak{u} \qquad (9)$$

$$\Delta T_{\underline{y}}=r_1Tc\mathfrak{u} \qquad (10)$$

hence

$$\frac{\Delta T}{T}=(p_1a+q_1b+r_1c)\mathfrak{u} \qquad (11)$$

If all dimensions are increased the same number of finite per-unit increments

$$a=b=c=1 \qquad (12)$$

and

$$\frac{\Delta T}{T}=(p_1+q_1+r_1)\mathfrak{u}=K\mathfrak{u} \qquad (13)$$

Thus, the simplicity of using infinitesimal per-unit increments becomes apparent. The per-unit increment in a dependent variable $A=BCD$ is the sum, not the complete complex product, of the individual per-unit increments. Writing the general total differential statement in a form particularly useful for this method

$$\frac{\Delta U}{U}=\frac{w}{U}\frac{\partial U}{\partial w}\mathfrak{u}_w+\frac{y}{U}\frac{\partial U}{\partial y}\mathfrak{u}_y+\frac{x}{U}\frac{\partial U}{\partial x}\mathfrak{u}_q\ldots \qquad (14)$$

At this point, an evaluation of the assumed shape is made on the basis of the following criterion. "If the device has optimum shape, the ratio between the resulting per-unit increments in optimizing characteristics is equal to the ratio between exponents or powers of linear dimension for these characteristics." See Appendix I for proof. For example, if volume responsive to the third power of linear dimension were being optimized against surface area responsive to the second power the application of increments to any one dimension should result in a 3-to-2 ratio of volume increments to surface increments for optimum shape, i.e., if shape is optimum and $(\Delta H)/H=m\mathfrak{u}$, the ratio which must hold is

$$\frac{\dfrac{\Delta V}{V_{\underline{H}}}}{\dfrac{\Delta S}{S_{\underline{H}}}}=\frac{3}{2} \qquad (15)$$

For $(\Delta D)/D=m\mathfrak{u}$

$$\frac{\dfrac{\Delta V}{V_{\underline{D}}}}{\dfrac{\Delta S}{S_{\underline{D}}}}=\frac{3}{2} \qquad (16)$$

If any one dimension is small in comparison to the other dimensions for optimum performance, applying increments to it will yield a greater than 3-to-2 ratio between volume and surface increments. The reverse is similarly true.

4. The fourth step is obvious. After evaluating the assumed shape, it can now be judiciously reproportioned toward the optimum. The amount of correction necessary is indicated by the extent to which the incremental characteristic ratios correspond to the desired linear dimension ratios. The increments are again applied and the second trial shape is evaluated.

5. In the fifth step, the process of correcting proportions and recalculation of rates of characteristic change is continued until the optimum shape is achieved. Few trial shapes are required if the work is pursued in a planned, orderly fashion. On some of the simpler problems, the application of infinitesimal per-unit increments leads directly to simple simultaneous algebraic equations which can be solved explicitly for the optimum shape. The real value of the method is not for solving this simple type of problem but rather for the more complex problem leading to simultaneous equations which can be solved only with great difficulty.

An Academic Illustrative Problem

The simple problem of proportioning a right circular cylinder, both ends closed, for the maximum volume and minimum surface will effectively illustrate the method. This is the practical problem in making economical tin cans from the material approach alone. It is a problem which is very easily solved by the mathematical method. From the mathematical solution the optimum proportions are realized when the diameter is equal to the height. Apply the method of infinitesimal per-unit increments as follows:

1. The expressions for volume and surface are

$$V=\frac{\pi D^2H}{4} \qquad (17)$$

$$S=\frac{\pi D^2}{2}+\pi DH \qquad (18)$$

2. Volume is responsive to the third

power of linear dimension and surface is responsive to the second power of linear dimension

$$V \propto l^3 \tag{19}$$

$$S \propto l^2 \tag{20}$$

3. Assume that the height is three-fourths of the diameter $H/D=3/4$. Apply an increase of three (or any number) infinitesimal per-unit increments to H and evaluate the resulting increments on V and S. They should have the ratio of 3 to 2. Regarding V, if $(\Delta H)/H=3\mathbf{u}$, then

$$\frac{\Delta V}{V_{\underline{H}}}=\frac{H}{V}\frac{\partial V}{\partial H}\frac{\Delta H}{H}=3\mathbf{u} \tag{21}$$

Regarding S, if $(\Delta H)/H=3\mathbf{u}$, then

$$\frac{\Delta S}{S_{\underline{H}}}=\frac{H}{S}\frac{\partial S}{\partial H}\frac{\Delta H}{H}=1.8\mathbf{u} \tag{22}$$

For this step it is convenient to evaluate $(\Delta V)/V_{\underline{H}}$ and $(\Delta S)/S_{\underline{H}}$ by simple inspection. It is obvious that V responds directly to the percentage change in H. S responds to the $(DH)/(D^2/2+DH)$ fraction of the percentage change of H evaluated on the basis of $H=3$, $D=4$, i.e.

$$\frac{\Delta S}{S}=\frac{12}{16/2+12}3\mathbf{u}=1.8\mathbf{u}$$

Evaluating

$$\frac{\dfrac{\Delta V}{V_{\underline{H}}}}{\dfrac{\Delta S}{S_{\underline{H}}}}=\frac{3\mathbf{u}}{1.8\mathbf{u}} \tag{23}$$

which is greater than the desired ratio of 3 to 2. Therefore H is too small with respect to D.

4. Since H is small with respect to D, increase H so that $H=D$. Regarding V, if $(\Delta H)/H=3\mathbf{u}$, then

$$\frac{\Delta V}{V_{\underline{H}}}=3\mathbf{u} \tag{24}$$

Regarding S, if $(\Delta H)/H=3\mathbf{u}$, then

$$\frac{\Delta S}{S_{\underline{H}}}=\frac{1}{1/2+1}3\mathbf{u}=2\mathbf{u} \tag{25}$$

Evaluating

$$\frac{\dfrac{\Delta V}{V_{\underline{H}}}}{\dfrac{\Delta S}{S_{\underline{H}}}}=\frac{3\mathbf{u}}{2\mathbf{u}} \tag{26}$$

as desired. Therefore the optimum shape of the right circular cylinder is realized for $H=D$.

Discussion of the Method

Any device can be described by a number of physical dimensions and other coefficients. A number of ratios, one less than the number of physical dimensions, completely establishes the configuration or shape of the device. If all of the dimensions of a device are increased by the same infinitesimal per-unit increment, the ratio between resulting increments in the characteristics for optimization will be equal to the ratio of the exponents of generalized linear dimension to which the characteristics are responsive. This is true for any shape. For the optimum shape, however, each individual dimension has the same degree of effect on characteristics as the others. Applying increments to one dimension alone under these circumstances approaches the same effectiveness as increasing all dimensions simultaneously and, hence, will also result in characteristic increments in proportion to the ratio of dimension exponents. See Appendix I for proof.

A significant factor in seeking optimums is seen by referring back to the cylinder problem. "Optimization is not achieved between two characteristics by holding one dimension constant and varying the other until the ratio of characteristics goes through a maximum." A maximum does not occur. V/S does not go through a maximum as H is increased for a fixed value of D. The optimum is truly defined only as the maximum value of volume V for a fixed value of surface S (or minimum value of S for fixed V). This is the condition sought out by the mathematical solution and the solution by infinitesimal per-unit increments.

The method is applicable to optimum determination where one ratio is restricted as well as the completely free case. This is of practical importance where it might be found that a completely free optimum yields ungainly proportions. By restricting one ratio to a fixed value, the optimum can be recalculated yielding a completely new optimum configuration.

It should be emphasized that some problems can very easily be solved by the generally useful mathematical method. It can consist of first taking the differential of the polynomials describing the two optimizing characteristics. These are set equal to zero for the condition of maximum-minimum and the ratios of coefficients of like differentials in the two expressions are equated. This yields simultaneous algebraic equations the solution of which represents optimum shape. This straightforward mathematical approach was used in solving the resistance-limited punched core reactors but became extremely difficult for the thermally limited reactor and the toroid cases.

The method described is a tool to use in conjunction with intimate knowledge of a component design. The designer will know whether his problem is yielding a maximum or a minimum from the nature of the optimizing characteristics. He will rarely find a practical problem where he has several maxima and even less rarely find a case where he can ambiguously solve for a minimum where a maximum is sought. Conventional mathematical means must otherwise be employed to answer this kind of question if the answers are not already evident in the character of the design.

Considerations for Optimizing Magnetic-Amplifier Reactors

Magnetic-amplifier reactors might be built as toroids using wound cores, or by conventional transformer techniques using stacked, punched cores. In small reactors, performance, gain, and time constant should be the criterion on which optimization is based. Economy of materials is comparatively unimportant since most of the cost for the small reactor is in assembly. Similarly, power efficiency is unimportant in considering operation costs. The most important item then is figure of merit, expressed as d-c power gain per second of control-circuit time constant. Gain should be optimized for a given time constant. This is true for both punched and wound cores.

For very large reactors (roughly 10 kw and above), economy of materials and power efficiency become important and performance assumes decreasing importance. It is desirable then to realize the greatest power output possible for a given cost or quantity of materials.

The foregoing represents one set of considerations for optimization. Particular applications might dictate optimizing characteristics other than these and will accordingly influence the assumptions, complexity of the analysis, and the resulting optimum shapes.

THE RESISTANCE-LIMITED OR SMALL REACTOR

Efficiency being unimportant in small reactors, it is possible to match the load resistance to the reactor winding resistance for maximum power transfer to the load. Under these circumstances, current density in the power winding is analytically related to the rating or size of the reactor such that, as rating increases, current density increases. This is developed in Appendix II along with the optimizing parameters. The maximum rating or size for which this relationship is valid occurs when current density is sufficiently high to tax, to the fullest, the power-dissipation ability of the reactor's surface. This defines the case for the resistance-limited reactor. From the analysis, the resistance-limited case is confined to reactors rated roughly 5 volt-amperes or smaller. At higher ratings, the reactor is governed by thermal limitations and load resistance must be larger than a-c winding resistance to keep current density within thermal allowances.

After-saturation characteristics of the reactor are negligible for this size of reactor, as experience shows that the maximum *H* or magnetizing force does not reach 1,000 oersteds. This is the point at which the air core flux of the winding introduces a noticeable addition to the relatively flat, intrinsic saturation flux density of the core material. The squareness of the hysteresis loop for the material affects the abruptness of firing but does not affect the substantially resistive character of load circuit impedance while load current flows.

The Thermally Limited or Large Reactor

Large reactors, ratings beyond the maximum of 5 volt-amperes for the resistance-limited case, are thermally limited. Current density in the a-c windings is analytically related to the amount of heat-dissipating surface the reactor possesses. This is developed in Appendix III. As the reactor is made larger, load resistance must be made a greater and greater proportion of the a-c circuit resistance to keep current density within thermal limitations. This improves reactor efficiency toward a limit of 100 per cent.

The analysis, to yield results, must be made on the basis of a negligible portion of input power being lost in the reactor, the ideal situation for the large reactor. Reactor current rating, on the other hand, is directly related to the reactor power loss. When a per-cent power loss for the reactor is chosen, consistent with what can be looked upon as practical negligible loss of input power, the reactor rating is at once defined analytically. 2.5 Per cent of input power was chosen as a reasonable figure and this yielded an output rating of about 10 kw. Since efficiency increases with rating, this means that the thermally limited analysis applied to ratings above 10 kw.

Optimizing Objective

With optimum shapes realizable for small reactors below 5 volt-amperes and large reactors above 10 kw, the undefined region in between represents an area of compromise between two kinds of optima. A semilogarithmic compromise between the shape ratios appears appropriate as a general guide for punched core reactors. As a matter of fact, these two power levels are figures safely in the power range of the aforementioned two general design cases. The analysis of a specific reactor arrangement will yield limiting power level between these. The analysis can be used to guide resistance-limited designs much larger than 5 volt-amperes and thermally limited designs much smaller than 10 kw, even to the point where these two limits overlap. Thus, the designer can weight his design in the 5-volt-amperes to 10-kw range in favor of the preferred of the two merit cases described.

Assumptions for the Analysis

Many factors influence the design of a reactor. This method permits the simultaneous consideration of more of these factors than is commonly convenient for the designer. The analysis covers the important factors. Assumptions are made regarding those which have little effect on optimization results. Considerations of these would yield very little improvement in the optimum but would add much to problem complexity. This does not imply that they are neglected in application considerations. It is assumed that:

1. Core losses can be neglected. For the 60-cycle resistance-limited reactor, core losses are such a small proportion of output power and reactor copper loss that they have only a small effect on shape. The thermally limited design is based on dissipating the greatest possible reactor losses without exceeding permissible reactor temperatures—a condition fulfilled in 60-cycle reactors when they are phased full on. Here, copper losses predominate and iron losses are very small since the core is fully or almost fully saturated during all parts of the cycle.

2. The cross section of the magnetic path is uniform and continuous. As long as the air gap of stacked cores is a small, fixed proportion of the magnetic path, the analysis is unaffected.

3. The rectifiers are ideal rectifiers. Silicon and germanium rectifiers with extremely high back resistance approximate this condition in reverse characteristics. This is important to the resistance-limited case where reverse current effects the gain portion of the optimizing characteristics as a function of shape. With selenium rectifiers, the optimum accuracy is not as good. On the other hand, in the thermally limited reactor analysis, rectifier leakage has negligible effect on reactor heating and maximum power output and, thus, has no practical effect on reactor shape. Rectifier forward voltage drop is a sufficiently small part of load circuit voltage so that it has little effect on shape determination. Furthermore, if the forward voltage drop were always a fixed proportion of the load circuit voltage, the analysis would be unaffected.

4. The after-saturation reactance of the reactor power winding can be neglected. It is responsive to shape changes in only a small way. The small values of *H* at which resistance-limited reactors operate make air core flux a much less significant factor than it is in even the thermally limited reactor. Finally, if after-saturation reactance as a significant quantity were always a fixed proportion of load circuit impedance, the analysis would be unaffected.

5. In resistance-limited reactors, power delivered to the load circuit is equal to power loss in the reactor winding. This is a reasonable, practical condition for small reactors.

6. Heat-dissipating surfaces dissipate heat at a uniform rate. By applying a uniform rate to only selected surfaces of the reactor, it is possible to approximate the condition in a practical reactor where the iron core is less effective in dissipating heat than the copper-winding surface.

Problem

To show in detail how the method was applied to a magnetic-amplifier problem, the details of optimizing the large thermally limited reactor are related in Appendix IV. This represents a problem that is extremely difficult to solve by regular mathematical techniques.

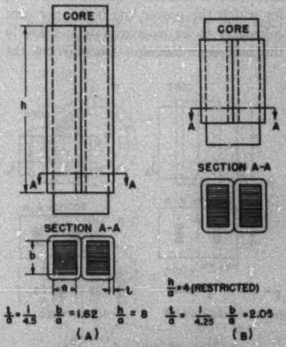

CORE
h
A
A
SECTION A-A
b
a
t
$\frac{t}{a}=\frac{1}{4.5}$ $\frac{b}{a}=1.62$ $\frac{h}{a}=8$
(A)
CORE
A
A
SECTION A-A
$\frac{h}{a}=4$ (RESTRICTED)
$\frac{t}{a}=\frac{1}{4.25}$ $\frac{b}{a}=2.05$
(B)

Fig. 1. Optimum core shape solutions for thermally limited reactors

A—Free optimum
B—Restricted optimum

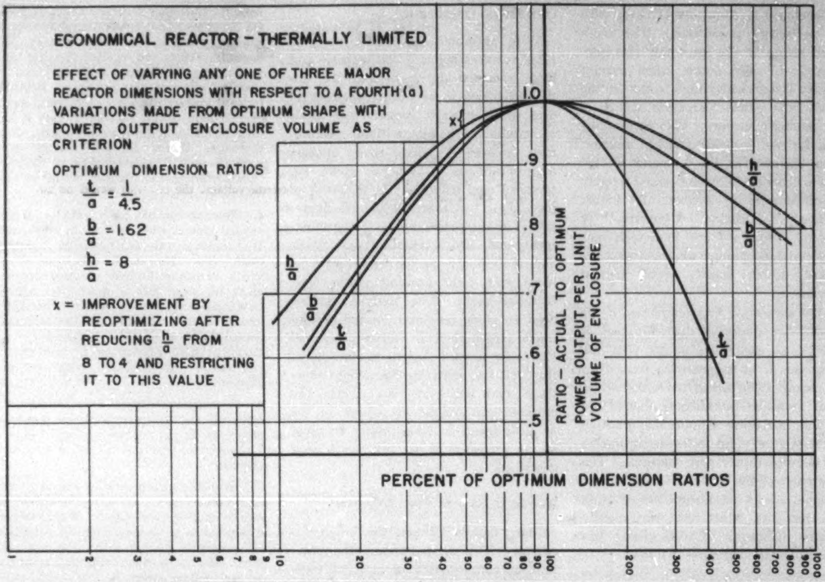


Fig. 2. Economical reactor, thermally limited

Presentation of General Magnetic-Amplifier Results

Thermally Limited Reactor— Maximum Output for Minimum Cost

From analysis in Appendix IV it is seen that the optimum shape for the reactor results in a very tall, elongated device. Fig. 1(A) shows the reactor shape drawn to scale. The reactor rating for which the power loss is 2.5 per cent calculates to be approximately 4 kw in the half-wave circuit. Considering the foregoing optimum unrealistic from the standpoint of proportions, the same problem can be solved restricting h/a to a value of 4 rather than the free value of 8. This is an arbitrary choice that is more appealing to the eyes and easier to build, mount, etc. The resulting configuration is shown in Fig. 1(B). These are proportions much more practical and the loss in optimum represents only 2.5 per cent of that achieved by the completely free solution.

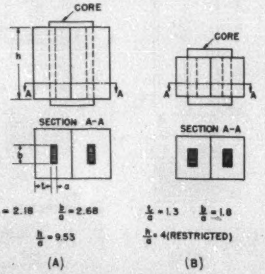


Fig. 3. Optimum core shape solutions for resistance-limited reactors

A—Free optimum
B—Restricted optimum

The optimum shown in Fig. 1 is not a sharply defined one. It is broad and flat. This is illustrated by the curves of Fig. 2 where the dimension ratios are varied one at a time from the full free optimum.

Resistance-Limited Reactors— Maximum Figure of Merit

An analysis of the resistance-limited reactor yielded a shape having a large window and thin leg. This shape is shown in Fig. 3(A). Maximum rating for this shape occurs when the size is large enough to encounter thermal limit. At 0.5 watt per square inch of surface, this is approximately a 6-watt output for the half-wave circuit.

Again the free optimum yields impractical proportions. By restricting h/a to a value of 4 rather than the free 9.53, a new optimum was obtained, as shown by Fig. 3(B). Maximum rating for this shape occurs at 3.5 watts half-wave. Optimum under this restriction has 8 per cent less merit than the free optimum. As in the thermally limited case, a broad flat optimum is seen to result for the resistance-limited case. This is shown in Fig. 4.

Continuing the resistance-limited analysis to toroids where the configuration of the winding is given an approximate trapezoidal form, an optimum is achieved such as shown in Fig. 5. Maximum rating for this reactor is approximately 3 watts.

Appendix I. Proof of the Method's Criterion

Let P and Q be the two characteristics that are important as design objectives in an optimization problem. Both P and Q are homogeneous polynomials and are appropriately described by the same dimensions, i.e., $l_1, l_2 \ldots l_n$

$$\begin{aligned} P &= P_1 + P_2 + \ldots P_s \\ &= (a_1 l_1^{\alpha_{11}} l_2^{\alpha_{21}} \ldots l_n^{\alpha_{n1}}) + (a_2 l_1^{\alpha_{12}} l_2^{\alpha_{22}} \ldots \\ &\qquad l_n^{\alpha_{n2}}) + \ldots (a_s l_1^{\alpha_{1s}} l_2^{\alpha_{2s}} \ldots l_n^{\alpha_{ns}}) \end{aligned} \qquad (27)$$

Vol.
Pt. 1
195

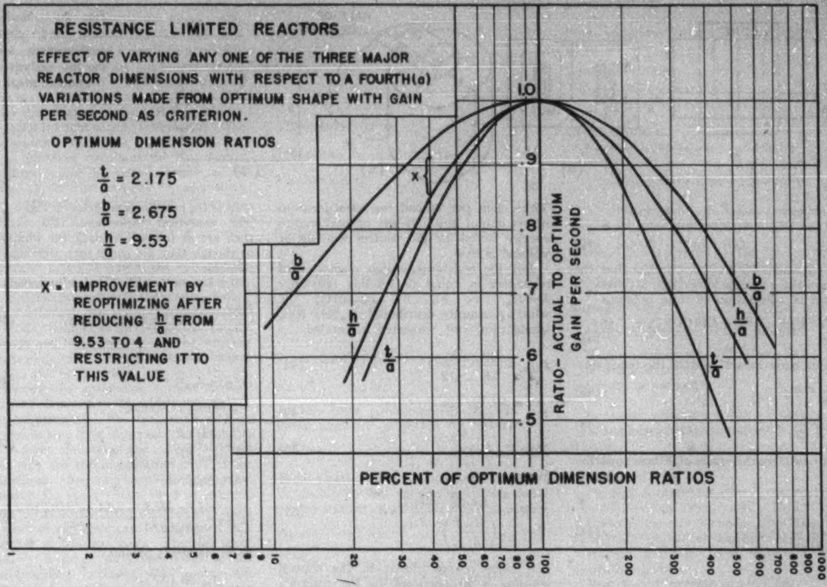


Fig. 4. Resistance-limited reactors

$$Q=Q_1+Q_2+Q_3\ldots Q_t=(b_1l_1^{\beta_{11}}l_2^{\beta_{21}}\ldots l_n^{\beta_{n1}})+(b_2l_1^{\beta_{12}}l_2^{\beta_{22}}\ldots l_n^{\beta_{n2}})+\ldots(b_tl_1^{\beta_{1t}}l_2^{\beta_{2t}}\ldots l_n^{\beta_{nt}}) \quad (28)$$

The first subscripts on the exponents indicate the particular dimension with which each is associated. The second subscript indicates the term of the polynomial in which it occurs. The polynomials being homogeneous

$$(\alpha_{11}+\alpha_{21}+\ldots\alpha_{n1})=(\alpha_{12}+\alpha_{22}+\ldots\alpha_{n2})=(\alpha_{1s}+a_{2s}+\ldots\alpha_{ns})=A \quad (29)$$

$$(\beta_{11}+\beta_{21}+\ldots\beta_{n1})=(\beta_{12}+\beta_{22}+\ldots\beta_{n2})=(\beta_{1t}+\beta_{2t}+\ldots\beta_{nt})=B \quad (30)$$

A is the power of linear dimension to which P is responsive. B is the power of linear dimension to which Q is responsive.

The criterion for the method states that at optimum shape application of infinitesimal per-unit increments to any one dimension l_p

$$\frac{\dfrac{\Delta P}{Pl_p}}{\dfrac{\Delta Q}{Ql_p}}=\frac{A}{B} \quad (31)$$

This can also be written as

$$\frac{\dfrac{1}{P}\dfrac{\partial P}{\partial l_p}\dfrac{dl_p}{l_p}l_p}{\dfrac{1}{Q}\dfrac{\partial Q}{\partial l_p}\dfrac{dl_p}{l_p}l_p}=\frac{A}{B}=\frac{\dfrac{1}{P}\dfrac{\partial P}{\partial l_p}}{\dfrac{1}{Q}\dfrac{\partial Q}{\partial l_p}} \quad (32)$$

To prove that equations 31 and 32 are true, solve the problem of maximizing P for a fixed value of Q. This will yield the same result if Q is minimized for a fixed value of P. Since Q is fixed

$$dQ=0=\sum_{p=1}^{p=n}\frac{\partial Q}{\partial l_p}dl_p \quad (33)$$

and since P is to be a maximum

$$dP=0=\sum_{p=1}^{p=n}\frac{\partial Q}{\partial l_p}dl_p \quad (34)$$

Thus, equations 33 and 34 are simultaneous equations of n variables of the form dl_p. For two such equations to be compatible, the ratios of the coefficients of like terms must all bear the same fixed value, i.e.

$$\frac{\dfrac{\partial P}{\partial l_p}}{\dfrac{\partial Q}{\partial l_p}}\equiv K \quad (35)$$

for any values of p from 1 to n.

Now apply this condition to the problem computing the partial derivatives from equations 27 and 28

$$\frac{\partial P}{\partial l_p}=\frac{\partial P_1}{\partial l_p}+\frac{\partial P_2}{\partial l_p}+\ldots\frac{\partial P_s}{\partial l_p} \quad (36)$$

From equation 27 it is apparent that this becomes

$$\frac{\partial P}{\partial l_p}=\frac{1}{l_p}(\alpha_{p1}P_1+\alpha_{p2}P_2+\ldots\alpha_{ps}P_s) \quad (37)$$

and similarly

$$\frac{\partial Q}{\partial l_p}=\frac{1}{l_p}(\beta_{p1}Q_1+\beta_{p2}Q_2+\ldots\beta_{pt}Q_t) \quad (38)$$

Substituting equations 37 and 38 into equation 35

$$\frac{1}{l_p}(\alpha_{p1}P_1+\alpha_{p2}P_2+\ldots\alpha_{ps}P_s)=K\frac{1}{l_p}(\beta_{p1}Q_1+\beta_{p2}Q_2+\ldots\beta_{pt}Q_t) \quad (39)$$

It is apparent that there are n such equations to be written, one for each value of p. If all n equations were added, the result is

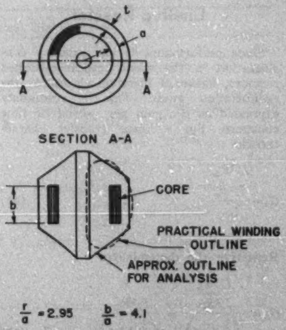


Fig. 5. Optimum core shape for resistance-limited toroidal reactor

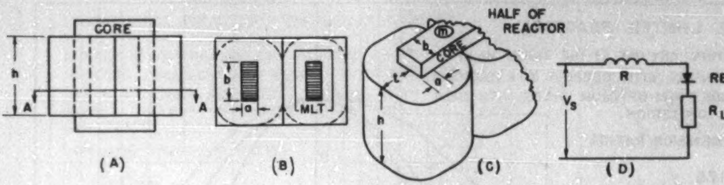


Fig. 6. Stacked core and circuit configurations used for the analysis of optimum shape

$$[(\alpha_{11}+\alpha_{12}+\ldots\alpha_{n1})P_1+\ldots(\alpha_{1s}+\alpha_{2s}+\ldots\alpha_{ns})P_s]=K[(\beta_{11}+\beta_{12}+\ldots\beta_{n1})Q_1+\ldots(\beta_{1t}+\beta_{2t}+\ldots\beta_{nt})Q_t] \quad (40)$$

But equations 29 and 30 stipulate that the sums of the α's and β's are equal to constants A and B. Equation 40 then becomes

$$A(P_1+P_2+\ldots P_s)=KB(Q_1+Q_2+\ldots Q_t) \quad (41)$$

From equations 27 and 28 this becomes

$$AP=KBQ \quad (42)$$

$$K=\frac{A}{B}\frac{P}{Q} \quad (43)$$

Substituting this value of K into equation 35

$$\frac{\frac{\partial P}{\partial l_p}}{\frac{\partial Q}{\partial l_p}}=\frac{A}{B}\frac{P}{Q} \quad (44)$$

or

$$\frac{\frac{1}{P}\frac{P}{\partial l_p}}{\frac{1}{Q}\frac{\partial Q}{\partial l_p}}=\frac{A}{B} \quad (45)$$

Thus, the method criterion as expressed by equations 31 and 32 is confirmed by the conventional solution of the problem on a generalized basis.

Appendix II. Optimization Parameters and Current Density Relationships for the Resistance-Limited Reactor

Since performance is the primary consideration in the small resistance-limited reactors, figure of merit is chosen as the optimization guide. This is commonly expressed as d-c gain per second of time constant. For a linear transfer characteristic

$$G_s=\frac{\frac{I_L^2R_L}{I_c^2R_c}}{\frac{L_c}{R_c}} \quad (46)$$

Rewriting the denominator

$$G_s=\frac{\frac{I_L^2R_L}{I_c^2R_c}}{2\left(\frac{\frac{1}{2}LI_c^2}{I_c^2R_c}\right)}=\frac{I_L^2R_L}{2\left(\frac{1}{2}LI_c^2\right)} \quad (47)$$

Thus, gain per second is proportional to power output and inversely proportional to energy stored in the control winding or volume of iron.

For the resistance-limited reactor, load resistance is equal to winding resistance $R=R_L$; see Fig. 6. Considering the windings equally distributed on both legs, single turn, and connected in parallel

$$I_{L\text{rms}}=\frac{V_{s\text{rms}}}{2R}\frac{1}{\sqrt{2}} \quad (48)$$

$$R=\frac{\rho(MLT)}{A_{ca}(SF)} \quad (49)$$

$$V_s=K_b(A_{fe}) \quad (50)$$

Hence

$$I_L^2R_L\propto\frac{(A_{fe})^2}{\left(\frac{MLT}{A_{ca}}\right)^2}\frac{MLT}{A_{ca}}\propto l^3 \quad (51)$$

and since energy stored in the control winding is proportional to the reactor's volume of iron

$$2\left(\frac{1}{2}LI_c^2\right)\propto l^3 \quad (52)$$

These are the powers to which the optimizing parameters are responsive.

Expressing load current and current density in terms of the physical dimensions

$$V_s=K_b(A_{fe})=K_bab \quad (53)$$

$$MLT=2a+2b+4t \quad (54)$$

$$A_{ca}=2th \quad (55)$$

$$I_{L\text{rms}}=\frac{K_b(SF)}{\rho\sqrt{2}}\frac{ab\,th}{2a+2b+4t}\propto l^3 \quad (56)$$

$$i_{a\text{rms}}=\frac{I_{L\text{rms}}}{A_{ca}}=\frac{K_b(SF)}{2\sqrt{2}\rho}\frac{ab}{2a+2b+4t} \quad (57)$$

Appendix III. Current Density Relationships in Large Thermally Limited Reactors

In the large thermally limited reactor with power winding equally distributed on both legs, it is recognized that the iron is less effective as a heat-dissipating surface than the exposed sides of the windings. This is true for heat generated in the windings because of the usual heavy insulation between the winding and core. The heat generated in the iron does not govern the design so it is neglected. To approximate this condition, all exposed surfaces of the windings are assumed to reject heat at 1/2 watt per square inch whereas the same is assumed for only the top and bottom ends of the punchings *m* circled in Fig. 6(C). The analytical expressions that evolve then are as follows, treating the windings as though they are single turn windings in parallel:

For constant watts per unit of surface

$$\frac{I_L^2R_w}{A_s}=K\leqq 0.5 \text{ watt per inches}^2 \quad (58)$$

Load current

$$I_L^2=(i_aA_{ca})^2 \quad (59)$$

Winding resistance

$$R_w=\frac{\rho(MLT)}{A_{ca}(SF)} \quad (60)$$

Winding loss

$$I_L^2R_w=\frac{I_L^2\rho(MLT)}{A_{ca}(SF)}=\frac{i_a^2A_{ca}\rho(MLT)}{(SF)} \quad (61)$$

From equations 58 and 61

$$I_L^2=\frac{A_{ca}KA_s(SF)}{\rho(MLT)} \quad (62)$$

also

$$i_a^2=\frac{KA_s(SF)}{A_{ca}\rho(MLT)}=\frac{K(SF)}{\rho}\frac{A_s}{Q_{ca}} \quad (63)$$

$$\frac{i_a^2Q_{ca}}{A_s}\frac{\rho}{(SF)}=K\propto\frac{i_a^2Q_{ca}}{A_s} \quad (64)$$

Reducing equations 62 and 63 to linear dimension relationships

$$I_L\propto l^{1.5} \quad (65)$$

$$i_a\propto l^{-0.5} \quad (66)$$

From reactor dimensions

$$MLT=2a+2b+\pi t \quad (67)$$

$$A_{ca}=2th \quad (68)$$

$$A_s=4ah+2hb+4\pi ht+4ab+8at+8bt+4\pi t^2 \quad (69)$$

Substituting these dimensional relations into equations 62 and 63

$$I_L^2=\frac{2K(SF)}{\rho}\frac{th}{2a+2b+\pi t}[4ah+2hb+4\pi ht+4ab+8at+8bt+4\pi t^2] \quad (70)$$

$$i_a^2=\frac{K(SF)}{\rho}\frac{[\ ]}{2th(2a+2b+\pi t)} \quad (71)$$

Appendix IV. Problem: Optimize the Large Thermally Limited Reactor

For simplicity of analysis the half-wave amplistat circuit of Fig. 6(D) will be used.

The reactor will be of a general configuration as shown in the same figure. The winding is equally divided between both legs, and is considered a single-turn winding and the two sections are treated as though connected in parallel. The control winding is made as small as practically possible under the design objective of this problem—the space it might occupy is neglected. The configuration need not digress from this for a full wave circuit as the two control windings would be simply connected in series. From the previous considerations, the optimum for the large thermally limited reactor should be based on the maximum power output for a given cost of materials.

Output rating is simply the volt-ampere product in the load. With the condition of negligible power loss in the power winding for economical operation, the output can be simply defined as the product of supply voltage and load current, i.e., the winding loss is neglected in computing power output but is the fundamental parameter in determining the maximum current density (surface dissipation) at which the reactor can be operated. The dissipation surface, as can be seen from Appendix III, includes all exposed surfaces of the copper windings and only the top and bottom ends of the punchings. Any surfaces deemed most realistic for the problem may be selected and will effect optimum shape accordingly.

Cost of materials can be expressed in a number of different ways. For this problem, it is taken as proportional to the parallelepiped enclosing the reactor—the product of the three maximum dimensions along the three major axes. This assumes that labor costs are directly proportional to material cost, copper has the same price per pound as the iron punchings, and that the extra volume of the parallelepiped represents the extra costs of mounting brackets, covers, and terminals. These are reasonable conditions for the large reactor but would not be so for the small reactors.

First develop the expressions for power output and cost volume: the power output

$$P_o \cong P_i = V_s I_L \qquad (72)$$

From the generated voltage expression

$$E_{rms} = \frac{2\pi}{\sqrt{2}} f N B_{max} A_{fe} \times 10^{-8} \text{ volts} \qquad (73)$$

it is apparent for this problem that f, N, and B_{max} are constant so that

$$V_s \propto A_{fe} = ab \qquad (74)$$

Current is established from equation 59

$$I_L = i_a A_{ca} \qquad (75)$$

Current density from equation 64

$$i_a = \left(\frac{K(SF)}{\rho} \frac{A_s}{Q_{ca}} \right)^{1/2} \qquad (76)$$

and in terms of reactor demensions, Fig. 6(C) and Appendix III

$$A_s = 4ah + 2hb + 4\pi ht + 4ab + 8at + 8bt + 4\pi t^2 \qquad (77)$$

$$Q_{ca} = A_{ca}(MLT) = 2th\,(2a + 2b + \pi t) \qquad (78)$$

Since

$$A_{ca} = 2th \qquad (79)$$

$$(MLT) = 2a + 2b + \pi t \qquad (80)$$

So from equations 72, 75, 76, and 79, power output

$$P_i \propto abthi_a \qquad (81)$$

Enclosure volume

$$\text{Height} = h + 2a \qquad (82)$$

$$\text{Width} = 2a + 4t \qquad (83)$$

$$\text{Depth} = b + 2t \qquad (84)$$

$$\text{Enclosure volume } Q_t = (h + 2a)(2a + 4t) \times (b + 2t) \qquad (85)$$

Second, establish the linear dimension relationships:
From equation 65

$$I_L \propto l^{1.5} \qquad (86)$$

From equation 74 it is apparent that

$$V_s \propto l^2 \qquad (87)$$

Then from equations 72, 86, and 87, power input (or output) is

$$P_i \propto l^{3.5} \qquad (88)$$

From equation 85 it is apparent that cost or parallelepiped volume is proportional to the product of three dimensions, hence

$$Q_t \propto l^3 \qquad (89)$$

Third, assume a shape (three ratios of the four dimensions) from whatever design history or experience is available and apply increments in turn to each individual dimension and see how closely the ratio of increments in power output and cost of enclosure volume follow the 3.5-to-3 law. When the correct shape is found, applying per-unit increments to any dimension will yield resulting per-unit increments in this ratio.

Assumed shape

$$a = 4t \qquad (90)$$

$$b = 2a = 8t \qquad (91)$$

$$h = 8a = 4b = 32t \qquad (92)$$

From equations 78 and 77 using equations 90 to 92

Winding volume

$$Q_{ca} = 2th(2a + 2b + \pi t) \qquad (93)$$
$$= 64t^2(8t + 16t + \pi t) = 1737t^3$$

Cooling surface

$$A_s = 4ah + 2hb + 4\pi ht + 4ab + 8at + 8bt + 4\pi t^2 \qquad (94)$$
$$= (512 + 492 + 128\pi + 128 + 32 + 64 + 4\pi)t^2$$

$$\text{Height} = h + 2a = (32 + 8)t = 40t \qquad (95)$$

$$\text{Width} = 2a + 4t = (8 + 4)t = 12t \qquad (96)$$

$$\text{Depth} = b + 2t = (8 + 12)t = 10t \qquad (97)$$

$$\text{Rating} \propto abthi_a \qquad (98)$$

Apply 1u increase to h dimension:

Height increase = 1×32/40 = 0.8u
Width unchanged = 0
Depth unchanged = 0

Enclosure volume increase = 0.8u

Rating increase then should = 0.8×3.5/3 = 0.933u.

Now winding volume increase from equation 93 = 1u.

Cooling surface increase from equation 94 = (512+492+128π)/1,642×1 = 0.856u.

Current density increase from equation 76 = (0.856−1)1/2 or decreases 0.072u.

Rating increase from equation 98 = 1−0.072 = 0.928u.

But desired increase was 0.933u.

The h increment increased rating 0.928/0.933 = 0.9464 times that desired. It is concluded that h is slightly too large with respect to other dimensions.

Apply 1u increase to b dimension by a sequence similar to the aforementioned. This b increment increased rating 0.9135/0.933 = 0.9791 times that desired. It is concluded that b is too large with respect to the other dimensions.

Apply 1u increase to t dimension by a sequence similar to the aforementioned. The t increment increased rating 0.579/

Table I. Per Unit of Desired Increase in Rating Resulting from Increments in h, b, t, and a

Trial	Dimension and Trial Equivalent							
	h	h_{eq}	b	b_{eq}	t	t_{eq}	a	a_{eq}
1	0.9464		0.9791	$h/4$	0.9325	$h/32$	1.0455	$h/8$
		$4b$				$b/8$		$b/2$
	h too large	$32t$	b too large	$8t$	t too large		a too small	$4t$
		$8a$		$2a$		$a/4$		
2	1.02		1.015	$h/4$	0.936	$h/28$	1.0154	$h/7$
		$4b$				$b/7$		$\frac{4}{7}b$
	h too small	$28t$	b too small	$7t$	t too large		a too small	$4t$
		$7a$		$\frac{7a}{4}$		$a/4$		
3	1.0397		1.00548	$\frac{7.3}{8(4.5)}h$	0.9765	$h/36$	1.0092	$h/8$
		$\frac{8(4.5)}{7.3}b$						$\frac{4.5}{7.3}b$
	h too small	$36t$	b too small	$7.3t$	t too large	$\frac{b}{(7.3)2}$	a too small	$4.5t$
		$8a$		$\frac{7.3}{4.5}a$		$a/4.5$		

0.621 = 0.9325 times that desired. It is concluded that *t* is too large with respect to the other dimensions.

Apply 1u increase to *a* dimension by a sequence similar to the aforementioned. The *a* increment increased rating 1.057/1.011 = 1.0455 times that desired. It is concluded that *a* is too small with respect to the other dimensions.

Fourth, the results of the first trial are tabulated in Table I along with the assumed ratios in the form of h_{eq} values. With the assumed ratios and resulting per-unit (of desired) increment thus logged for ready review, the next trial can be easily formulated and evaluated. For a new assumed shape, similar to equations 90 to 92, it is necessary to prepare another group of equations similar to equations 93 through 98.

Fifth, the process is repeated and, finally after a few trial, the resulting per-unit increase in rating is sufficiently close to the desired one so that the last assumed shape can be accepted as the answer.

Reference

1. MAGNETIC CIRCUITS AND TRANSFORMERS (book), **EE Staff, M.I.T.** John Wiley and Sons, Inc., New York, N. Y., 1952, chap. VIII, article 5.

Determination of Thermal Life of Enameled Wire by Laboratory Methods

F. A. SATTLER
NONMEMBER AIEE

A THERMAL life test which will give quick results is a desirable feature in wire enamel development. Since it may be necessary to screen many formulations, it is desirable that this test be convenient to run. Four criteria for thermal life determination are briefly reviewed. These are as follows: 1. visual inspection of thermally aged coils; 2. flexibility of aged enameled wire; 3. decrease in film thickness with aging; and 4. dielectric breakdown of aged dielectric twist samples. In each test the enameled wire samples are aged in air-circulating ovens and the life is determined at temperatures of 100 to 300 degrees centigrade (C). Some data obtained from these tests are presented.

Visual Inspection Test

In this test a 15-inch length of enameled wire is wrapped around a 5/8-inch mandrel, removed from the mandrel, and identified with an aluminum tag. The 5/8-inch mandrel is convenient to handle and is large enough to eliminate heat-shock failure for most materials. Heat-shock failure is the cracking of an enamel film caused by a combination of film strain and short-time heating. The resulting coil is shown in Fig. 1.

Each coil is aged at a specific temperature in an open glass container in an air-circulating oven along with other samples. The coils are examined periodically at 22.5×magnification. Failure is evidenced by general crazing, cracking, or flaking. A typical crazing pattern is shown in Fig. 2.

When failure occurs by flaking, the end point is not as sharp and is more difficult to determine. The failure point from flaking has arbitrarily been set as the time at which five large flaws or ten small flaws are seen on the sample.

In Fig. 3, a plot of log life versus reciprocal of absolute temperature[1] is given for Formvar, enamel *A* (organic enamel) and enamel *B* (modified silicone), all on no. 18 American Wire Gauge copper wire. The failure points fall along a straight line for most materials at temperatures below 200 C, but the slope generally falls off at the higher temperatures.

Wire sizes nos. 20 to 15 American Wire Gauge have given fairly reproducible results in this test. The small wire sizes, no. 26 American Wire Gauge and smaller, have been observed to show an appreciably longer life than the larger wire with the same enamel. There is no clear-cut explanation for this phenomenon at the present time. Frequency of observation (thermal cycling) and coil diameter have been observed to have little effect on the thermal life results in this test.

This method can also be used to study impregnating varnish-wire enamel combinations. The enameled wire coils are dipped in the varnishes and baked out before aging. In Fig. 4, note that Formvar appears to be unaided thermally by a silicone impregnating varnish, but is improved appreciably by a phenolic impregnating varnish. Enamel *A*, on the other hand, develops thermal cracks more readily with a phenolic varnish impregnant than does the unimpregnated enameled wire, while a silicone impregnant improves thermal life by a factor of four.

The visual inspection test has also been used in testing sealed, air-excluded systems such as enameled wire in transformer oil, or in Freon-oil mixtures. Fig. 5 shows two samples of enameled wire sealed in vacuo in transformer oil in a glass tube. A sample of Formvar in this test is flawless after 495 days at 175 C, while an air-exposed sample of Formvar crazes in about 30 days at 175 C.

The visual inspection test may give misleading results, especially at temperatures above 200 C. Some enamel films retain a good appearance while evaporating rapidly, leaving very little insulation. Another disadvantage is the personal element involved in determining the end point. On the other hand, the test is very easily set up and readily shows up large differences in thermal stabilities.

Flexibility Test

While the low temperature data are considered the most reliable and most im-

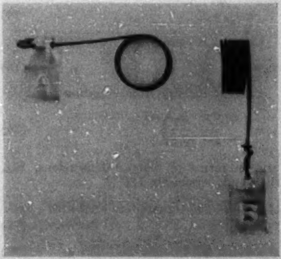


Fig. 1. Visual inspection coil

Paper 55-47, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing December 7, 1954.

F. A. SATTLER is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

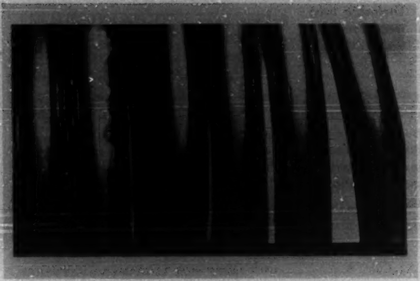
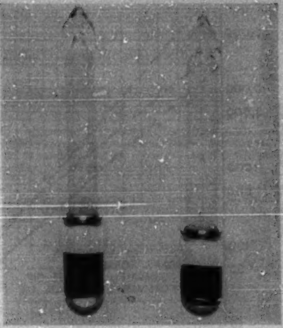


Fig. 2. Formvar 2 months at 175 C

Fig. 5 (right). Enameled wires sealed in transformer oil 495 days at 175 C



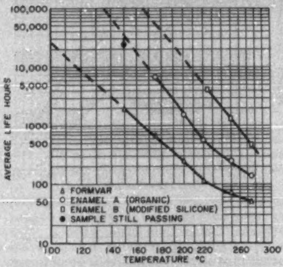


Fig. 3. Visual inspection test. Average life in hours versus temperature in C (1/T scale), no. 18 American Wire Gauge enameled wires

Fig. 6 (right). Percent elongation to failure (no. 18 American Wire Gauge) versus aging time at 200 C

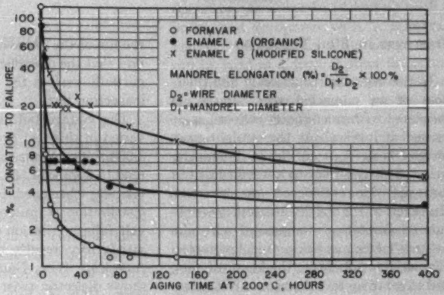
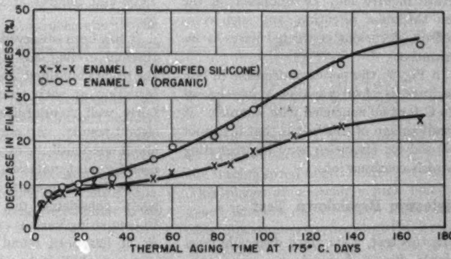


Fig. 7 (right). Percent decrease in film thickness of enameled wire versus aging time at 175 C, days



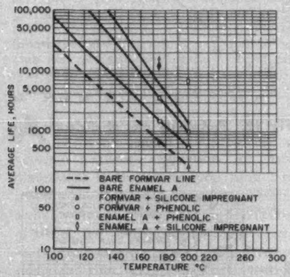


Fig. 4. Visual inspection test. Average life in hours versus temperature in C, no. 18 American Wire Gauge enameled wires

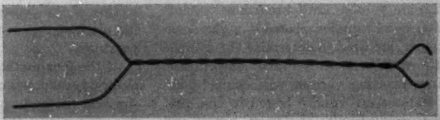


Fig. 8 (right). Dielectric twist sample

portant in the previously described static life test, it is not always expedient to await such data. Thus a flex-life test may show the trend of degradation of the film long before it is realized in a static life test.

In this flexibility test, straight pieces of enameled wire are aged in open glass tubes in ovens at temperatures in the range of 100 to 200 C. The flexibility retention is determined periodically by winding on graduated mandrels. In Fig. 6, per-cent elongation to failure is plotted against aging time at 200 C for several enamels. It is difficult to interpret the data in terms of expected life, but some background is being obtained on this test. The method cannot be used in studying wire enamel impregnating varnish combinations.

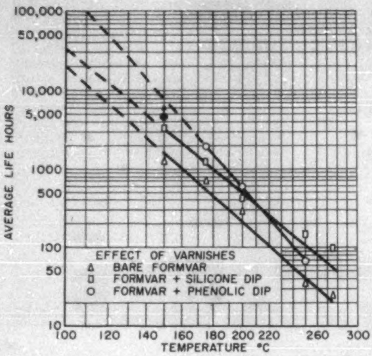


Fig. 9 (left). Dielectric twist life in hours versus temperature in C, 500-volt criterion

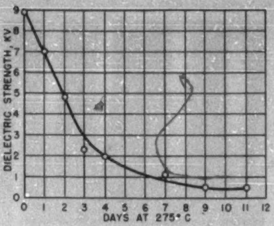


Fig. 10 (right). Dielectric strength in kilovolts versus days at 275 C. Enamel B silicone impregnated

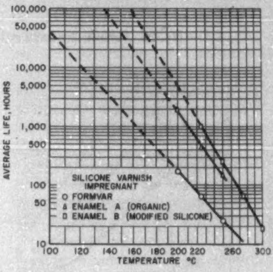


Fig. 11 (right). Dielectric twist life in hours (1,000 volts) versus temperature in C

Decrease in Film Thickness

In this test, the reduction in film thickness of an enameled wire sample, as measured with micrometer calipers, is recorded at intervals as the straight piece of wire is aged at an elevated temperature. This is similar to a weight loss test and may be considered less accurate. However, in the weight loss test it is difficult to simulate the enameled wire film in such factors as baking out low molecular weight polymers, thickness of film, and effect of metal substrate. With careful caliper measurements, a fairly accurate picture may be obtained on the film thickness retention, and such comparison of various enameled wires is informative.

In Fig. 7, the per-cent decrease in film thickness is plotted against aging time at 175 C for two enameled wire samples. A disadvantage of this test is that it cannot be used for enameled wire impregnating varnish combinations.

Dielectric Breakdown Test

In this test, a standard dielectric twist sample, Fig. 8, is aged at a given temperature and tested periodically at a specific voltage until breakdown occurs. Ten to 25 samples, placed in a glass container, are used to determine the average life of an enamel at each temperature. Impregnating varnish wire enamel combinations are readily tested in this procedure by dipping the twist sample and baking. In Fig. 9, the life-temperature curve is shown for Formvar enameled wire of no. 18 American Wire Gauge with no impregnation, silicone varnish impregnation, and phenolic varnish impregnation. It should be noted that the silicone varnish impregnation improves the life of the Formvar but not to the extent that the phenolic varnish does. This partially verifies the results of the visual inspection test.

The voltage failure criterion has been varied in this test over a range of 500 to 3,000 volts. Some enameled wire samples tend to approach 500 volts almost asymptotically when aged, Fig. 10, giving a wide distribution of failure points. A 1,000-volt failure criterion has been used with satisfactory results on nos. 17 and 18 American Wire Gauge wires. Fig. 11 shows dielectric twist life in hours versus temperature. A distribution curve for 1,000-volt failures at 250 C is shown in Fig. 12.

It has been observed that frequency of thermal cycling has a definite effect on the results obtained, and frequency of observation at each temperature must be fairly well standardized to obtain consistent results. An observation period of about one-tenth the expected life has been used for most materials.

The low temperature failure data are again considered the most important. In comparing various experimental enamels it has been found advantageous to plot actual dielectric breakdown values versus time at 200 C to obtain a quicker picture of the degree of degradation taking place at this comparatively low temperature.

Of the four methods described, the dielectric breakdown criterion is considered the best method of thermal life evaluation. However, the other three methods help to complete the picture and allow more definite conclusions to be drawn.

A comparison of the thermal life data of Formvar enameled wire as determined by visual inspection, dielectric breakdown, and accelerated motor testing[2] is shown in Fig. 13. The motor data curve

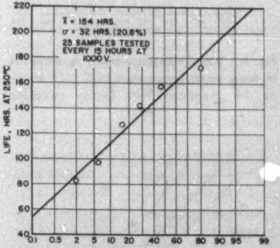


Fig. 12. Probability plot of dielectric twist failures

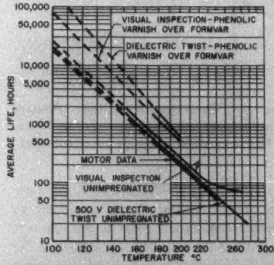


Fig. 13. Average life in hours versus temperature in C. Comparison of Formvar life estimated in various tests

Vol
Pt.
195

is low, especially as the Formvar enameled wire in the motor was impregnated with a phenolic varnish. However, the motor test involved severe vibration and humidity cycling superimposed upon thermal degradation.

Conclusions

The four criteria discussed for evaluating thermal stability of enameled wire, namely, visual inspection, flexibility, film thickness measurement, and dielectric breakdown, give a picture of the rate and type of degradation taking place when enameled wire is aged thermally. Data from the visual inspection test and dielectric breakdown test can be plotted on log life versus reciprocal of absolute temperature graphs to estimate practical operating temperatures of the enameled wires. Results from these tests show the same temperature dependence as motor life tests. Impregnating varnish wire enamel combinations can also be evaluated using the visual inspection and dielectric breakdown methods.

References

1. ELECTRICAL INSULATION DETERIORATION TREATED AS A CHEMICAL RATE PHENOMENON, T. W. Dakin. *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 113–18.

2. A METHOD OF EVALUATING INSULATION SYSTEMS IN MOTORS, C. B. Leape, J. McDonald, G. P. Gibson. *AIEE Transactions*, vol. 72, pt. III, Aug. 1953, pp. 793–99.

An Automatic Field Plotter

R. GELFAND
ASSOCIATE MEMBER AIEE

B. J. SHINN
STUDENT MEMBER AIEE

F. B. TUTEUR
ASSOCIATE MEMBER AIEE

Synopsis: Solutions of field problems can often be approximately obtained by constructing a 2-dimensional model of the field to be studied in an electrolytic tank. A servomechanism is described which is arranged to follow the equipotential lines that exist in such a model and to plot them automatically. A rotating co-ordinate system is employed to enable the device to trace out closed curves and other multivalued functions. This co-ordinate system is automatically aligned with the equipotential lines to be followed so that the sweep velocity is constant and parallel to the desired curve and correction velocity is normal to this curve. The system is analyzed with particular emphasis on stability and accuracy. The prototype plotter was able to follow curves with a minimum radius of $^1/_2$ inch and maintain an accuracy of 0.5 per cent of maximum travel.

AN analogue method commonly used in the solution of field problems involving Laplace's equation in two dimensions is the electrolytic tank. In its simplest form such a tank consists of a shallow basin of large area in the center of which the boundaries of the problem to be studied are placed in the form of electrodes that are connected to a fixed voltage source. If the area of the tank is sufficiently large, it is possible to ignore the edge effects introduced by the finite extent of the conducting medium. The solution to the problem is obtained in the form of equipotential lines which may be determined by direct measurement, and which may then be plotted. When a large number of points are to be located, this process is apt to become rather tedious, and automatic plotting becomes desirable.

Several improvements in the basic technique have been made in recent years. An improved arrangement of the tank itself was described by Boothroyd, Cherry, and Makar.[1] Their tank consists essentially of two conducting circular surfaces, placed back to back separated by an insulator and joined at the circumference. It can be shown that this configuration represents a transformation of an infinite conducting plane into a finite area; hence, it permits accurate analogues to be set up in small tanks with only negligible edge effects. An automatic plotter was described by Green,[2] but it suffers from the disadvantage that it cannot plot multivalued functions easily. The device described in this paper overcomes this disadvantage, and incorporates a variable gain feature that permits dependable operation over wide variations of potential gradient. J. K. Mickelson[3] described a device which handles multivalued functions and appears to be very similar to the one described here.

Elementary Theory of Operation

Assume that it has been possible to set up an electric analogue of a particular field problem by means of a plane of resistive conducting material with the proper boundaries and electrodes. With such an arrangement, the lines of constant voltage correspond to equipotential lines of the problem. The solution can be obtained by plotting selected equipotential lines e.g., all lines with integral values of voltage in the region of interest. The system constructed used an electrolytic tank as the conducting surface; however, the techniques to be described are equally applicable to any conducting surface.

The voltage at any position of the operating plane can be picked up with a probe and compared with a reference voltage. The voltage difference or error can be used in a servo system to correct the probe position to the desired line on the plane. For example, let the rectangular co-ordinates x and y be assigned to the surface of the tank and assume the equipotential lines are single-valued in y. If a constant velocity is applied in the x direction and the servo loop has freedom in the y direction, the probe will trace out the desired equipotential line.

With a constant sweep speed, the velocity of the probe along the curve will not in general be constant but will vary with the slope of the curve. The velocity can be found as follows

$ds=\sqrt{(dx)^2+(dy)^2}$ = infinitesimal distance along the curve

$$\dot{s}=\sqrt{\dot{x}^2+\left(\frac{dy}{dx}\dot{x}\right)^2}=\dot{x}\sqrt{1+\left(\frac{dy}{dx}\right)^2} \qquad (1)$$

From equation 1, it can be seen that, as the curve becomes parallel to the y axis, the tracing speed must become very large. This places an extremely high response speed demand on the servo loop. It is also important to notice that this system is incapable of plotting curves multivalued in y. A method will now be

---

Paper 55-104, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted May 26, 1954; made available for printing December 8, 1954.

R. GELFAND is with the University of Pennsylvania, Philadelphia, Pa., B. J. SHINN is with the General Electric Company, Schenectady, N. Y., and F. B. TUTEUR is with Yale University, New Haven, Conn.

The authors wish to acknowledge the assistance of Prof. P. M. Schultheiss who helped in the solution of many of the design problems encountered in this development. Particular thanks are also due to S. C. Bogan who first suggested the method of continuously varying the alignment of the auxiliary co-ordinate system.

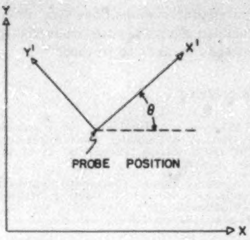


Fig. 1. Rotated co-ordinate system

described which will overcome these limitations.

Consider a new rectangular co-ordinate system having $x'-y'$ axes centered at the probe and rotated so that the x' axis makes an angle θ with the x axis; see Fig. 1. Let the sweep velocity be applied in the x' direction and let the servo system operate in the y' direction. Supposing that the angle θ between the two co-ordinate systems is manually adjustable, then a human operator watching the progress of the probe through the electrolyte would be able to keep always the x' axis, and thus the sweep direction, parallel to the equipotential line being followed. In this way, the tracing velocity could be maintained constant and there would be no difficulty in the following of multivalued lines since all lines would be single-valued in the neighborhood of the probe.

In practice it is difficult to rotate the co-ordinate axes physically, and it is therefore desirable to control the probes along a fixed set of axes, yet retain the flexibility of the rotating co-ordinate system. This end may be achieved by the use of two resolvers connected as shown in Fig. 2. The operation of these resolvers is, perhaps, best understood by assuming, for the time being, that the error voltage V_e and the angle θ are both zero. Under these conditions, the sweep voltage is applied to the x axis drive motor only and the probe moves parallel to the x axis at sweep speed. If the input shaft of the resolver is now rotated in the positive direction, the x axis motor receives a smaller voltage $V_s \cos \theta$ and the y axis motor receives the voltage $V_s \sin \theta$. The result is that the probe moves in the new direction θ but still with the same sweep speed. If the error voltage is also applied, the probe acquires a component of velocity perpendicular to the sweep velocity. The way in which all four velocity vectors add to give the final probe velocity is shown in Fig. 3, where S_x and S_y represent the sweep velocity components and C_x and C_y the correction velocity components in the x and y direction respectively.

The adjustment of the angle θ, supposed to require a human operator, may, of course, be made automatically. If an auxiliary probe is mounted near the main probe described previously, a voltage will exist between these two probes unless they are both located on the same equipotential line. This voltage may be used as the input to a servomechanism which will rotate the two probes so that they are always on the same equipotential line and, at the same time, vary the shaft angle of the resolvers. The adjustment of this angle will not, in general, be completely exact, particularly in regions where the equipotential lines are sharply curved; however, by spacing the two probes very closely together ($^1/_2$ inch or less), satisfactory performance may be obtained.

Description of the Experimental System

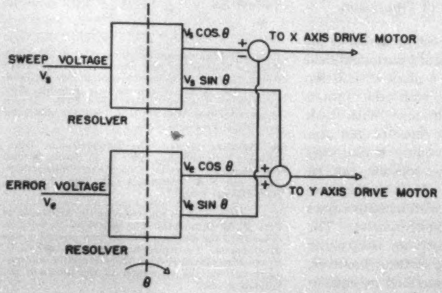
The two probes described must be continuously aligned and positioned in an electrolytic tank used as the resistive plane. The alignment is accomplished by means of a drive motor geared to a polystyrene rotor in which the probes are mounted; probe voltages are picked off by slip rings. A selsyn attached to the rotor shaft is used to sense the angular alignment of the probes. This entire assembly is mounted on a carriage, Fig. 4, which rides on a crossarm. The carriage is positioned by a cable from the y axis drive motor, while another cable from the x axis motor moves the crossarm along the two parallel supporting rails, see Fig. 5.

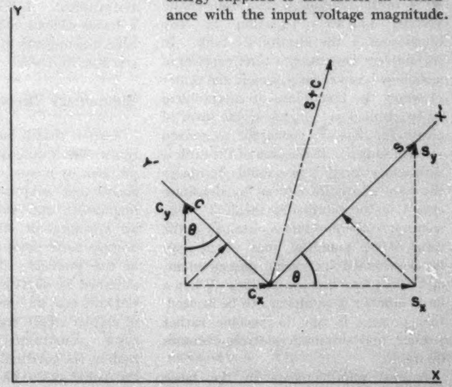
The resolvers required to transform the sweep and error voltages into the appropriate x and y components may either be mounted on the carriage and geared directly to the probe, or they may be mounted externally and driven by an electromechanical follow-up loop. The latter method was chosen so as to keep carriage size and weight down to a minimum. A selsyn, geared to the input shaft of the resolvers, is used with the selsyn coupled to the probe-turning motor in a conventional servomechanism which positions the resolvers in accordance with the probe angle.

The drive motors are 27-volt high-speed units with built-in gear boxes giving an output no-load speed of 250 rpm. These units were removed from a war-surplus *B-29* gun computer, as were the relay power amplifiers used to drive them. The relay amplifiers have an internal feedback arrangement which modulates the pulse width and repetition rate of the energy supplied to the motor in accordance with the input voltage magnitude.

Fig. 2 (below). Resolver connections

Fig. 3 (right). Velocity components





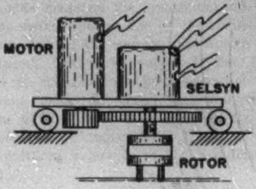


Fig. 4. Carriage assembly

Fig. 5. Carriage, crossarm, supporting rails, and electrolytic tank

The output-versus-input curve was found to be essentially linear between a dead-band region and saturation. The system loop gain is proportioned so that the dead-band referred to the pickup probes gives an error within the desired accuracy limits.

The resolvers are selsyns, also removed from the gunsight computer. Orthogonal trigonometric functions are obtained from the output by means of a resistive network; see Appendix I. The two resolvers, probe position selsyn, and resolver-control drive motor are mounted on a single chassis and coupled together by gears.

A-c amplifiers, cathode followers, and discriminators round out the components required. High-gain d-c amplifiers were avoided by using a 400-cycle carrier system and demodulating at a high-voltage level just ahead of the equalizer networks which feed the power amplifiers. The electrolytic tank used was modeled after the unit described by Boothroyd, Cherry, and Makar[1] to minimize edge effects.

Design Considerations

A block diagram of the complete system is shown in Fig. 6. The system has two input adjustments; the reference voltage V_r which selects the desired equipotential line, and the sweep voltage V_s which determines the sweep speed of the probe. Four different motors must be controlled: the two that move the carriage along the x and y axes, the motor which regulates the angle θ between the rotating $x'y'$ co-ordinate system and the fixed reference, and the follow-up motor that rotates the resolvers. Since the servomechanism controlling the angle θ is in many ways the heart of the system, it will be considered first. This servomechanism will be called the θ loop.

As shown in Fig. 7, the correct angle of alignment of the x' axis, ϕ, is the angle between the tangent to the desired curve and the x axis. This angle is determined approximately by the two probes referred to earlier. If the angle between a line drawn through the two probes and the x axis is called θ, then on the assumption of a uniform potential gradient the voltage between the probes is

$$e_\theta = KE\sin(\phi-\theta) \qquad (2)$$

where E is the voltage gradient, and K is a proportionality constant which depends on probe spacing and the dimensional units used. The difference between ϕ and θ is small if the θ loop operates properly so that equation 2 may be written approximately as

$$e_\theta \cong KE(\phi-\theta) \qquad (3)$$

If the gain of the amplifier in the θ loop is G_θ, then, in Laplace transform notation

$$\theta(s) = G_\theta H_\theta(s) e_\theta(s) \qquad (4)$$

where $H_\theta(s)$ stands for the frequency-sensitive elements in the θ drive system and represents equalizing networks, filters, and the drive motor. The negative of the loop gain of the θ loop is therefore

$$A_\theta(s) = KE\,G_\theta H_\theta(s) \qquad (5)$$

Note that the potential gradient E enters directly into this expression; this point will be considered in more detail later.

The θ loop is a standard unity-ratio feedback loop which takes the input ϕ and delivers the output θ. Hence, the well-known result

$$\theta(s) = \frac{A_\theta(s)\phi(s)}{1+A_\theta(s)} = \frac{KE\,G_\theta H_\theta(s)\phi(s)}{1+KE\,G_\theta H_\theta(s)} \qquad (6)$$

The follow-up loop that is required to rotate the resolvers is entirely conventional and requires no special comment. In the subsequent analysis it is assumed to be a perfect servo, i.e., the shaft position of the resolvers is always identically equal to θ the probe angle.

The remaining loop is required to move the probe designated as the "main probe" to the desired equipotential line. The voltage picked up by this probe is subtracted from the reference voltage V_r, and the difference $e_{y'}$ represents the error in main probe position. This voltage will be directly proportional to the perpendicular distance between the desired line and the main probe—assuming constant gradient, see Fig. 8. Thus, the error voltage will be $e_{y'} = E\Delta n$ where Δn is the perpendicular distance from the desired voltage line to the main probe. If $\Delta y'$ is the position error in the y' loop,

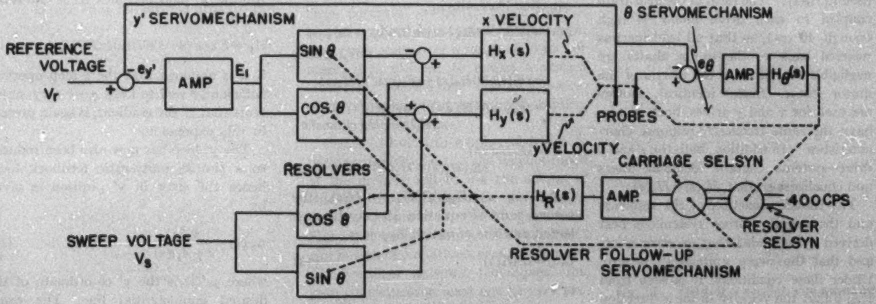


Fig. 6. System block diagram

1, 74
1
55

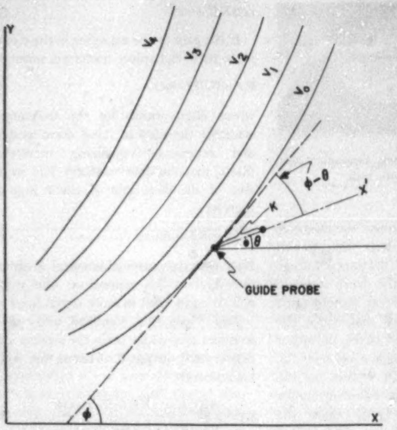
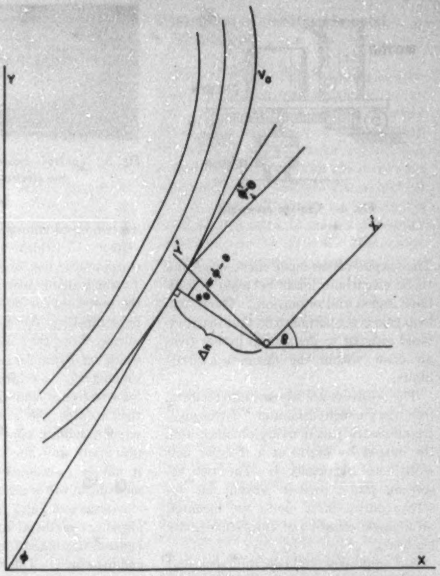


Fig. 7 (above). Correct angle of alignment of the x′ axis

Fig. 8 (right). Effect of misalignment of x′ axis



Vol.
Pt. 1
1955

i.e., the distance between the main probe and the desired equipotential line as measured along the y' axis, then

$$\Delta n = \Delta y' \cos(\phi-\theta) \qquad (7)$$

and

$$e_{y'} = E\cos(\phi-\theta)\Delta y' \qquad (8)$$

Note that the error voltage in the y' loop is a function of the error in the θ loop; if $\phi-\theta$ is greater than $\pm\pi/2$ radians, the y' loop becomes unstable. Hence, the error in the θ loop must be kept small for proper operation of the y' loop.

The error voltage $e_{y'}$ is amplified, passed through the resolvers, and is then applied to $H_x(s)$ and $H_y(s)$ which represent the frequency-sensitive element of the x and y drive systems similar to those used in $H_\theta(s)$. The ratio of the gear train coupled to each drive motor is high enough, 40 to 1, so that all load inertias referred back to the motor shafts are negligible compared to the inertia of the motor rotor. Since identical motors are used for x and y drives, both motors have the same frequency-response characteristics. In addition, both the x and y drive systems contain identical filters and equalizers so that $H_x(s)=H_y(s)$.

The transfer function of the resolvers and the x and y drive systems is best derived by assuming that an error exists and that the sweep voltage V_s is zero. Under these conditions, it is seen from Fig. 6 that the velocity in the x direction is

$$C_x(s) = -sE_1H_x(s)\sin\theta \qquad (9)$$

and the velocity in the y direction is

$$C_y(s) = sE_1H_y(s)\cos\theta \qquad (10)$$

The extra s in these expressions is required to make the definitions for $H_x(s)$ and $H_y(s)$ consistent with that already adopted for $H_\theta(s)$. Reference to Fig. 3 indicates that the components in the y' axis of C_x and C_y are $-C_x\sin\theta$ and $C_y\cos\theta$ respectively so that

$$C(s) = sy'(s) = -C_x(s)\sin\theta + C_y(s)\cos\theta \qquad (11)$$

Substituting equation 9 and 10 into equation 11 results in

$$\begin{aligned} C(s) = sy'(s) &= sE_1(-\sin\theta)H_x(s)(-\sin\theta) + sE_1(\cos\theta)H_y(s)\cos\theta \\ &= sE_1(\sin^2\theta)H_x(s) + sE_1(\cos^2\theta)H_y(s) \\ &= sE_1\{\sin^2\theta H_x(s) + \cos^2\theta H_x(s) + [H_y(s)-H_x(s)]\cos^2\theta\} \\ &= sE_1H_x(s) + sE_1[H_y(s)-H_x(s)]\cos^2\theta \end{aligned} \qquad (12)$$

A long as $H_x(s) \cong H_y(s) = H(s)$, the second term of equation 12 may be neglected and the equation becomes

$$C(s) = sE_1H(s) \qquad (13)$$

or

$$y'(s) = E_1H(s) \qquad (14)$$

Since E_1 is proportional to the error in probe position, the velocity of the probe in a direction normal to the desired line is proportional to the magnitude of the error. It can, incidentally, be shown in a similar manner that the tangential velocity of the probe can be expressed by

$$S(s) = sV_sH(s) \qquad (15)$$

Thus the sweep and error-correcting functions take place independently even though both drive motors handle these two functions simultaneously.

If the gain of the amplifier in the y' loop is equal to $G=E_1/e_{y'}$, the loop gain for the y' loop becomes from equations 8 and 14

$$A_{y'} = E\cos(\phi-\theta)GH(s) \cong EGH(s) \qquad (16)$$

if it is assumed that the θ loop operates sufficiently well to keep $\phi-\theta$ very small. Note that E, the gradient, is again present in this expression.

The y' loop has now also been reduced to a simple, unity-ratio feedback loop; hence the error in y' position is given by

$$\Delta_{y'}(s) = \frac{y_i'(s)}{1+A_{y'}(s)} \qquad (17)$$

where $y_i'(s)$ is the y' co-ordinate of the desired equipotential line. The exact nature of $y_i'(s)$ and, therefore, the demands made on the system depend on the sweep velocity and on the radius of curvature of the equipotential line to be traced. Let it be supposed that the sweep speed is fixed by specification at some value S_0. Then the expression for $y_i'(s)$ is derived by assuming the y' loop to be inoperative. Under these conditions, reference to Fig. 8 indicates that

$$y_i'(t) = S_0 \int \tan(\phi - \theta)dt$$
$$\cong S_0 \int (\phi - \theta)dt \qquad (18)$$

if $\phi - \theta$ is small. Taking the Laplace transform of both sides results in

$$y_i'(s) = \frac{S_0[\phi(s) - \theta(s)]}{s} \qquad (19)$$

which by use of equation 6 may be written as

$$y_i'(s) = \frac{S_0\phi(s)}{s[1 + KEG_\theta H_\theta(s)]} \qquad (20)$$

Substitution of equations 20 and 16 in equation 17 results in

$$\Delta y'(s) = \frac{S_0\phi(s)}{s[1 + KEG_\theta H_\theta(s)][1 + EGH(s)]} \qquad (21)$$

To illustrate a rough design procedure, all filter and equalizer time constants in $H_\theta(s)$ and $H(s)$ are neglected and only the integration inherent in the drive motors is retained. Thus

$$KG_\theta H_\theta(s) = \frac{K_1}{s}$$
$$GH(s) = \frac{K_2}{s} \qquad (22)$$

Then

$$\Delta y'(s) = \frac{S_0 s\phi(s)}{(s + K_1E)(s + K_2E)}$$
$$= \frac{S_0\phi(s)}{(s + K_1E)(s + K_2E)} \qquad (23)$$

It is seen that the error is directly proportional to the sweep speed S_0 and the rate of change of the slope of the equipotential line ϕ; hence, there is theoretically no error in following a straight line. A more severe test is the tracing of a small circle of radius R, for which ϕ is equal to S_0/R.

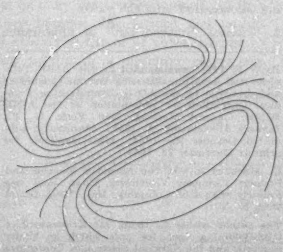


Fig. 9. Field of parallel plane conductors

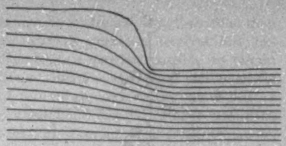


Fig. 10. Field between a plane and stepped conductor

Application of the final value theorem[4] to equation 23 shows that the final error resulting in this test is

$$\Delta y' = \frac{S_0^2}{RK_1K_2E^2} \qquad (24)$$

Thus, a minimum value of $K_1K_2E^2$ is established if specifications of tracing velocity, minimum radius of curvature, and maximum allowable error are given. For the device described here, these specifications were 2 inches per second, $^1/_2$ inch, and 0.1 inch respectively.

Effect of Potential Gradient

The equations indicate clearly the effect of the potential gradient on the accuracy of the system; it is seen that the error is inversely proportional to the square of the gradient. It has also been shown that the loop gains of both the θ and y' loops are directly proportional to the gradient; see equations 5 and 16. Although the simple loop gain functions assumed in equation 22 would represent stable systems no matter what the gain is, in practice the effect of additional time lags in the loops will tend to cause the system to become unstable in regions of large potential gradient. Since the potential gradient in a typical configuration may easily vary over a range of 10 to 1 or more, it is difficult to realize a system combining satisfactory accuracy in regions of low gradient with good stability in high-gradient areas.

One solution is to employ a device to measure the gradient in the neighborhood of the probes, and to use the output of this device to vary the gain of the amplifiers by means of a variable-gain tube. Since loop gain variations of about 2 to 1 can be handled relatively easily with standard equalizer network techniques, this gain compensation need not be very precise. A gradient detector may be realized in the following way: Assume that the θ loop is operating properly and the two probes are aligned on an equipotential line. Let one additional probe be placed at right angles to a line passing through the original two probes so that the three probes form an isosceles right triangle with the main probe at the apex. The voltage between the main probe and the new probe is then a measure of the potential gradient and may be applied to the second grid of the variable μ tube. This technique results in a highly stable system which maintains good accuracy even in regions of relatively low potential gradient.

Recording

The plotting of an equipotential configuration involves not only the tracing of equipotential lines but also the recording of them. The use of the electrolytic tank in this particular system eliminated the possibility of recording directly on the working plane. Thus, voltages proportional to probe position were obtained from 10-turn helipots mechanically coupled to the x and y drive motors. If these voltages are fed to an oscilloscope with d-c amplifiers in the y and x axes, the beam will trace out the equipotentials. An oscilloscope camera with the shutter left open for the duration of the tracing of each equipotential will give a recording of the desired plot.

The voltages could also be fed to an automatic plotting table such as the variplotter. This method was the one used to obtain the three sample plots shown in Figs. 9, 10, and 11. Fig. 9 represents a parallel plane capacitor or a parallel strip transmission line. The field between a plane conductor and a stepped conductor is shown in Fig. 10. Fig. 11 shows the well-known voltage equipotentials for a parallel wire transmission line.

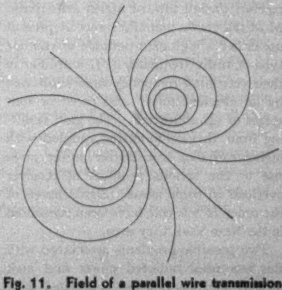


Fig. 11. Field of a parallel wire transmission line

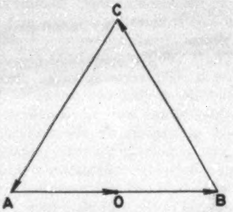


Fig. 12. Space vector diagram of a selsyn output

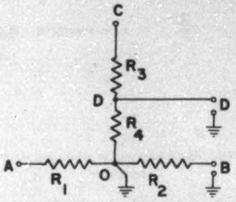


Fig. 13. Resistive network to obtain orthogonal trigonometric functions from a selsyn

Conclusion

The automatic field plotter offers a means of rapid and accurate determination of electric field configurations. Such a device should prove especially useful in situations requiring a large number of successive plots where parameters are varied to obtain an optimum configuration.

Appendix I. The Resolvers

It is well known that, if the rotor of a selsyn (with a nonsalient rotor) is excited with a carrier frequency ω, the amplitude of the carrier voltage between any two terminals will vary sinusoidally with rotor-angular rotation. Thus, the voltage between terminals A and B could be represented by

$$V_{AB}=(GV_R\sin\theta)(\sin\omega t)$$

where θ is angle of rotation, V_R is the peak amplitude of rotor voltage, and G is a constant depending on turns ratio, etc. The other two output voltages attain a maximum amplitude at angles of rotation of 120 degrees more and 120 degrees less than the angle at which V_{AB} is maximum

$$V_{BC}=GV_R\sin(\theta+120^\circ)\sin\omega t$$

and

$$V_{CA}=GV_R\sin(\theta-120^\circ)\sin\omega t$$

Note that the phase of the carrier frequency is the same for all three voltages.

A space vector diagram may be set up to represent the amplitude modulation factors $\sin\theta$, $\sin(\theta+120^\circ)$, and $\sin(\theta-120^\circ)$. This is shown in Fig. 12.

If two equal resistors are connected between points A and B, then the voltages across them are $V_{AO}=V_{OB}={}^1/_2 V_{AB}$. The vector representation of the voltage from their center to point C is seen to be

$$V_{OC}=1.732\,V_{OB}\underline{/+90^\circ}$$

The voltage V_{OC} may be attenuated to equal V_{OB} in magnitude by a resistance attenuator; See Fig. 13.

As long as the total resistance of this attenuator is large compared to the other resistors, no loading will result and V_{AO} will still be in phase with V_{OB}. Thus, the voltage V_{OD} will be 90 degrees out of space phase with V_{OB} and

$$V_{OD}=GV_R\cos\theta\sin\omega t$$

$$V_{OC}=GV_R\sin\theta\sin\omega t$$

It is apparent that the selsyn plus resistive network described are the equivalent of a resolver which produces orthogonal trigonometric functions from an input consisting of angular rotation of the rotor and electric excitation of the stator.

References

1. An Electrolytic Tank for the Measurement of Steady State Response, **A. R. Boothroyd, E. C. Cherry, R. Makar.** *Journal*, Institution of Electrical Engineers, London, England, vol. 96^1, 1949, pp. 163–77.

2. Automatic Plotting of Electric Fields, **Paul E. Green.** *Review of Scientific Instruments*, New York, N. Y., vol. 19, no. 10, Oct. 1948, pp, 646–53.

3. Automatic Equipment and Techniques for Field Mapping, **J. K. Mickelson.** *General Electric Review*, Schenectady, N. Y., vol. 52, no. 11, Nov. 1949, p. 19.

4. Principles of Servomechanisms, **G. S. Brown, D. P. Campbell.** John Wiley and Sons, Inc., New York, N. Y., 1948, p. 72.

Television Receiver Signal Overload

C. MASUCCI
ASSOCIATE MEMBER AIEE

THE PROBLEM of signal overload has become increasingly severe as a result of the trend in modern television receiver design toward higher sensitivity. Most television manufacturers at present use either a high-gain pentode or cascode type of radio frequency (r-f) amplifier in their very-high-frequency tuners followed by an efficient mixer and three or four stages of amplification. The gain resulting from this combination coupled with higher radiated transmitter power gives rise to the serious problem of receiver overload in some areas. Signal levels of the order of 1.0 volt have been measured in the New York City area.

Two possible problems associated with this increased radiated power and high receiver sensitivities are concerned with the prevention of receiver overload in areas where all signals are high, and the prevention of receiver overload in areas where some signals are high and some are weak.

This latter problem requires for best design and consumer appeal an automatic device which will reduce the receiver's sensitivity or gain for the strong channel and restore it on the weak channel. This paper will discuss this method and also compromise methods for an economy design. The same designs can apply equally well to the first overload condition, i.e., where all signals are of high levels.

Overload Manifestations

Signal overloading in a television receiver manifests itself in numerous ways. A few examples follow.

1. Cross-modulation interference between channels not necessarily adjacent; and between television channels and frequency modulation stations; etc.

2. Noisy pictures in high-signal areas. This is a result of operating r-f and intermediate frequency (i-f) tubes near or beyond cutoff.

3. Sixty-cycle synchronizing buzz especially in intercarrier receivers where the sound is taken from the plate of the video amplifier.

4. Loss of picture synchronization as a result of the compression of the synchronizing information into the video.

5. Picture "washed out" in appearance

Paper 55-139, recommended by the AIEE Television and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing December 23, 1954.

C. Masucci is with the American Machine and Foundry Company, Greenwich, Conn., and was formerly with CBS-Columbia, Inc., Long Island City, N. Y.

The author wishes to thank M. Kronenberg of CBS-Columbia, Inc., for data which he kindly supplied, and the National Electronics Conference of 1953 for use of portions of the paper "An Automatic Television Overload Elimination Circuit."[1]

Vol.
Pt. 1
195

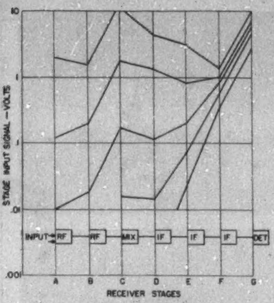


Fig. 1. **Receiver signal distribution for constant antenna input levels**

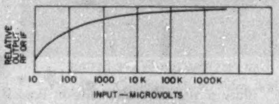


Fig. 2. **Typical stage performance for increasing input signal**

with improper operation of the contrast control.

6. Improper operation of the automatic gain control (AGC) circuit.

7. Complete reversal of picture with, of course, very poor synchronization.

Fig. 1 is a graphic presentation of the signal voltages at various stages for fixed antenna input levels. As can be seen from the signal distribution throughout the receiver, there may be more than one point of overload as the input signal increases. Also, for each individual stage, the signal output (Fig. 2) increases up to a given level and then remains relatively constant for any additional input signal. The point of "flattening" can be called the overload level for each stage.

Another factor of importance to the problem of overload, in addition to the higher gain receivers and higher radiated transmitter power, is that some receivers can be more susceptible to overload than others because of variations, such as cutoff characteristics of i-f tubes, r-f and i-f gain, maximum or minimum power supply voltage available, etc. These variations, of course, are expected and they cannot be controlled any closer than they are without increasing receiver cost, except that it may be possible to obtain closer control of the cutoff characteristic of the tubes used in the controlled i-f stages. Some companies have used tube selection to permit all receivers to handle, at least up to a point, the same signal variations, e.g., 20 to 100,000 microvolts with suitable AGC systems. In addition, tube manufacturers have recently announced the development of tubes with more uniform or remote cutoff characteristics to replace the standard *6CB6* as a controlled i-f stage in a television receiver. They are the miniature types *6CF6* and *6DC6*.

System Considerations

A desirable overload system should supply the necessary attenuation for strong signals in such a manner as to prevent grid rectification in the first r-f stage. The attenuation should be removed when the signal decreases to a predetermined level. It is essential, however, that insertion of attenuation, if automatic, be at a signal level other than that at which it is removed, so that a given signal will not cause the system to "chatter" or "motor-boat."

The possible methods of reducing the probability of receiver overload can be divided into two main groups:

1. Methods of eliminating signal overload in existing television receivers:
   a. Insertion of a resistance attenuator into the antenna lead-in.
   b. Insertion of a resistance attenuator into the r-f tuner section.
   c. Reduction of antenna gain.
   d. Use of quarter-wave stubs.
2. Methods of eliminating signal overload by correctly designed television receiver:
   a. Automatic electronic control: use of a broad-base amplified AGC system.
   b. Automatic electromechanical control: use of a relay to insert a step into the AGC control characteristic.
   c. Manual electromechanical control: use of a switch or potentiometer to insert a step into the AGC control characteristic.

Circuit Considerations

Some of the methods of eliminating signal overload in existing television receivers will be treated first.

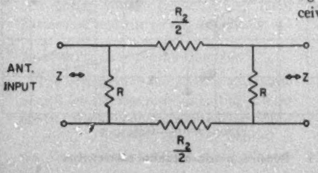


Fig. 3 (left). **A symmetrical π attenuation pad**

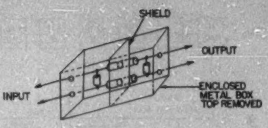


Fig. 4. **Constructional features of the π attenuation pad**

Resistance Attenuation in the Antenna Lead-In

An economical and effective attenuator pad to use is the symmetrical π type shown in Fig. 3.

The unknown values R_1 and R_2 can be determined by the equations

$$R_2 = Z\frac{K^2-1}{2K}$$

$$R_1 = Z\frac{K+1}{K-1}$$

where

Z = antenna and receiver input impedance
K = the voltage or current ratio of attenuation from input to output

Thus for an antenna and receiver input impedance of 300 ohms and a current ratio of attenuation equal to 10 $R_1 = 366$ ohms and $R_2 = 1{,}485$ ohms.

Suitable standard Radio-Electronic-Television Manufacturers Association values for R_1 and $R_{2/2}$ would be 390 and 750 ohms respectively. These choices result in the pad appearing as approximately 325 ohms instead of 300, which is not a sufficient mismatch to the antenna and receiver terminals to cause trouble.

For most effective results the pad should be constructed as sketched in Fig. 4. This method of construction is suggested since the stray signal feedthrough from input to output is reduced to just the distributed capacitance of $R_{2/2}$ which is quite small. The full attenuation for which the pad was designed can therefore be realized. In addition, the pads should be placed as close to the tuner input as practical to eliminate r-f pickup from tuner to antenna input.

This method of preventing receiver signal overload will, of course, result in a reduction of all signals whether strong or weak. Care must therefore be exercised in the amount of attenuation used to prevent unnecessary deterioration of weak signals to favor strong signals.

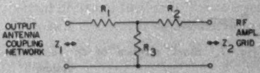


Fig. 5 (right). **An unbalanced T-attenuation pad**

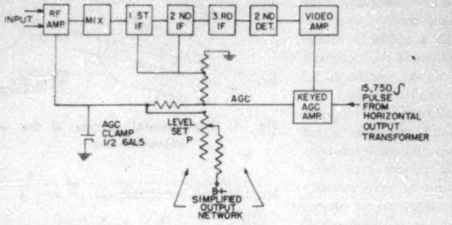


Fig. 6 (left). Block diagram of a keyed AGC type of control system (an automatic electronic control system)

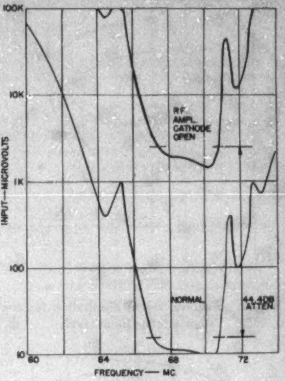


Fig. 8 (right). Comparison of a receiver over-all response with and without overload control

Resistance Attenuation in the R-F Tuner Section

The T attenuator pad shown schematically in Fig. 5 can be used where it is desired to insert attenuation directly into the r-f tuner section. The unknown values R_1, R_2, and R_3 can be determined by using the equations

$$R_1 = Z_1 \frac{(N+1)}{(N-1)} - R_3$$

$$R_2 = Z_2 \frac{(N+1)}{(N-1)} - R_3$$

$$R_3 = 2\, \frac{\sqrt{N Z_1 Z_2}}{N-1}$$

where

Z_1 = output impedance of antenna matching transformer

Z_2 = input impedance of r-f amplifier tube grid

N = power ratio of attenuation from input to output (always greater than unity)

The anticipated attenuation may not be achieved because of the difficulty of shielding input from output when the three resistors have to be inserted on the r-f portion of a turret tuner strip. In fact, it is probable that attenuations greater than 20 decibels may not be feasible. But 20 decibels attenuation will cure many of the overload problems. This method, of course, permits weak signals in the same area on other channels to be received with maximum receiver performance. For the two methods described, there are commercially available economical pads which will give the desired results in most instances.

Reduction of Antenna Gain

The antenna system may have too much gain for the high-signal area in question. A stacked antenna may have been used where only an ordinary dipole with reflector would be sufficient. Of course, the antenna choice must be consistent with good reception on strong and weak signals, if such are available.

If cost is no object, two or more antennas could be used and their gain selected according to the strength of the signal received. The consumer could select the antenna for the particular channel he desires to watch by means of a good quality low-capacity switch. The emphasis on a good switch is necessary so that the intercoupling between antennas is at a minimum. It is also necessary, for low intercoupling, that the lead-ins be spaced as far apart as practical and only come together at the switch. It is also possible by means of commercially available rotatable antennas to reduce the signal input to a receiver, consistent with good reception, and eliminate signal overload.

Use of Quarter-Wave Stubs

Since a resonant transmission line, if opened at one end, presents a short circuit at the first 1/4-wave point, it can be used to act as a short circuit at the picture carrier frequency of the channel being overloaded. A perfect short circuit would attenuate the channel completely. Therefore, the short circuit should be good enough to eliminate signal overload but not entirely lose the picture or sound.

The equation for computing the length of 1/4-wave stub in inches for a 300-ohm twin lead is

$$L = \frac{2{,}425}{f}$$

where f = frequency in megacycles.

A stub for channel 2 which has a video carrier frequency 55.25 megacycles would have a length $L = 2{,}425/55.25$.

This is the value for a theoretically perfect short circuit. The Q of the stub line can be lowered by termination with a

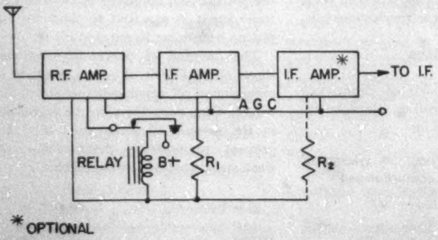


Fig. 7. Block diagram of an automatic electromechanical control system

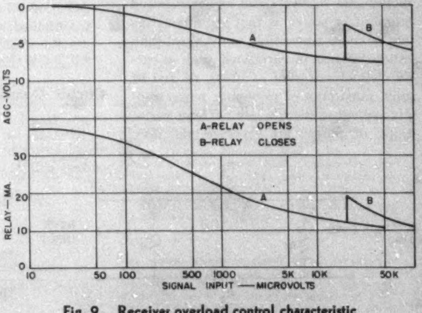


Fig. 9. Receiver overload control characteristic

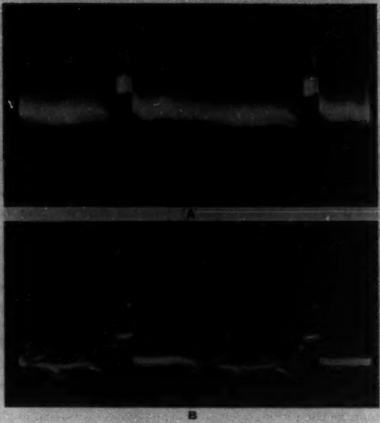


Fig. 10. Oscillograph of video and sync signal

A—Under overload conditions. B—With elimination of overload conditions

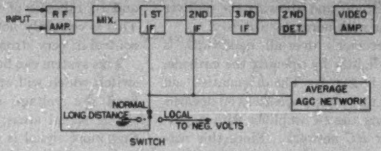


Fig. 11. Block diagram of a manual electromechanical control system

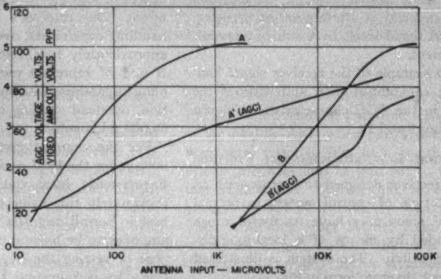


Fig. 12. Comparison of receiver control characteristic and output signal with and without overload

suitable resistance to maintain the correct frequency response within the channel. A method often used in the field consists of shortening the line by cutting 1/8-inch segments until the desired picture attenuation with good sound is attained. Since a stub out for one channel may resonate, e.g. in the 3/4-wave length mode at a higher channel it is necessary that reception on all other channels be carefully checked for picture quality. The stub should not be draped around the antenna lead-in but should for best results be taped to the cabinet if it is of wood. If the cabinet is metal, the stub should be taped to the fiber back of the receiver. Reception of all other channels should be checked to insure good quality.

Methods

The remainder of the paper will discuss methods that can be used in the design of a receiver to permit it to handle signals of high amplitude. In all the methods to be described, the final effect is to insert into the AGC characteristic a step which will permit the receiver to accept larger amplitude signals before the development of negative voltage (AGC action) again flattens off and the receiver overloads.

AUTOMATIC ELECTRONIC CONTROL

The system to be discussed is automatic except for the necessity of setting a level control for best operation in a particular location. However, once the control is set, the operation of the receiver is automatic even though the range of signals available may vary from the very weak to the very strong.

A keyed AGC type of control system, as illustrated in Fig. 6, has inherently a wide range of operation. In particular, the negative voltage developed by such a system is of the order of 10 or more volts. This is a result of: 1. the system operating on sync tips due to the horizontal pulse keying action, and 2. d-c gain.

The system may develop sufficient negative voltage under strong signal conditions to cut off the r-f and i-f tubes. The resultant noisy picture is a by-product of a poor tuner noise factor. To offset this effect, the "level-set" potentiometer *P* as shown in Fig. 6, is adjusted to vary the delay in the action of the AGC on the r-f amplifier during weak signal conditions. The adjustment made for a particular receiver location ensures good reception under strong signal conditions and has little effect on moderate signals.

Since the potentiometer is connected through a large resistor to the power source, it is possible under weak-signal conditions for the effective AGC voltage to be positive. To prevent operating the r-f and i-f tube with a positive voltage on the grids, an AGC clamp 1/2 *6AL5* or equivalent, is used to hold the positive voltage to the minimum, i.e., the drop across a conductive diode. With these refinements, the system can and does perform satisfactorily under most signal conditions.

AUTOMATIC ELECTROMECHANICAL CONTROL

For this method of control, attenuation is inserted into the r-f circuit by means of a relay which is electronically controlled to put a step into the AGC control characteristic of the receiver. Fig. 7 is a block diagram of the method and indicates that it requires the addition of one relay and two resistors (R_1 and R_2) to increase the signal handling capabilities of most present-day receivers.

Essentially, at the time the receiver line switch is thrown, as the receiver warms up, the current flow in the stages will close the relay and it will be held closed by the added r-f amplifier current. As the signal strength increases, the AGC voltage becomes more negative and relay coil current decreases. As the overload point is approached, the decreasing relay current allows the contacts to open the r-f amplifier cathode, thus attenuating the signal and reducing the current flow through the relay coil. The AGC voltage decreases, thereby increasing the current through the i-f stages and relay coil but not sufficiently to close the relay. If the signal decreases further or the receiver is switched to a weaker signal, the decrease in AGC supplies sufficient current to actuate the relay, closing the r-f amplifier cathode circuit and allowing the receiver to operate with normal sensitivity.

The receiver's over-all bandwidth is affected slightly by opening the cathode, as shown in Fig. 8. The attenuation can be varied up to approximately 50 decibels by bridging the r-f amplifier with a suitable *R* or *RC* network. Since this network can remain in the circuit at all times, it has the distinct advantage of requiring no manual switching. The system is sufficiently flexible to permit the insertion and removal of attenuation at predetermined signal levels by a simple relay adjustment.

An example of the receiver signal handling capabilities with and without the step in the AGC characteristic is shown in Figs. 9 and 10.

Manual Electromechanical Control

A receiver designed to operate with an AGC type of control system in normal signal areas may have its useful range extended by means of a local-long-distance switch. The system is illustrated in block diagram form in Fig. 11.

In this method of signal overload control, the AGC voltage is obtained from the second detector and is of necessity a voltage which will vary with the picture content. Also, the maximum negative voltage available for controlling the r-f and i-f tubes is of the order of 5 volts. This is certainly not sufficient to permit control in very strong signal areas.

This system can be aided by the use of a switch which will apply a more negative fixed d-c voltage to the amplifier, in strong signal areas. The result will be that more signal is required to start the average AGC action and therefore the signal-handling ability of the receiver will be extended. This is illustrated in the curves of Fig. 12. Curves *A* and *A'* represent the receiver's normal signal-handling capabilities, overload starting at approximately 1,500 microvolts. Curves *B* and *B'* represent receiver signal-handling capabilities with switch in local position, overload starting at approximately 50,000 microvolts.

For the example given, the negative voltage for the local position was 9 volts. Experiments have indicated that approximately this value of voltage is the best to permit any *3BC5* used as the r-f amplifier to be beyond cutoff. The purpose of setting the negative voltage at this value is to increase the probability of the local-position performance to be acceptable in many receivers and locations.

This method of control can be improved by the addition of a level setter, that is, a potentiometer to vary the negative voltage applied when the switch is in the local position. This will permit the receiver to accept even stronger signals than those indicated in Fig. 12, while maintaining an acceptable over-all noise factor. The tuner noise factor can be further improved by an additional switch section which will reduce the mixer gain along with the r-f amplifier gain when the receiver is operating in the local position. The third position of the local-normal-long-distance switch is useful to obtain the maximum possible receiver gain to permit acceptable fringe area reception.

The methods described and illustrated have all been used with satisfactory results, either alone or in combination. The method or methods used for a particular case will depend on the severity of the problem and whether or not cost is an important factor.

References

1. An Automatic Television Overload Elimination Circuit, C. Masucci, J. R. Peltz, W. B. Whalley. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 9, 1953, pp. 575–79.

2. Tube Selection Increases Signal Capacities of TV Receivers, R. G. Horner. *Tele-Tech and Electronic Industries*, New York, N. Y., vol. 13, no. 10, 1954, pp. 70–71; 132–33.

3. Automatic Gain Control for Television Receivers, K. R. Wendt, A. C. Schroeder. *Television, RCA Review*, Princeton, N. J., vol. 5, pp. 149–61.

4. A New Fast Noise-Immune Television AGC Circuit, K. R. Wendt. *Ibid.*, pp. 161–69.

Vol. 7
Pt. 1
1955

Science, Statistics, and Reliability Engineering

J. C. BEAR
NONMEMBER AIEE

SCIENCE may be defined as "systematized knowledge." Statistics has been called "the technology of the scientific method."[1] Engineering is "the useful application of the findings of science." At first glance, it would appear that these three fields are bound together in an inseparable relationship; but we know that they have always been treated as quite separate areas of interest, both on the campus and in industrial application, and therein lies a problem. While in certain types of engineering the theories of science may be applied without very much knowledge of how these theories were derived, there are other fields of engineering that must use the method of science itself. Reliability engineering is an example. It has arisen with the multiplying complexity and importance of electronic products, and must cope with all of the phases in the evolution of a product, from the designer to the user, and back to the designer. The recent rise of such fields as systems engineering, operations research, and cybernetics seems to have been motivated by this same increasing complexity. Engineers at present are faced with problems of a more and more confusing nature, the solutions for which lie in a more and more theoretical realm. We believe that the traditional, scholastic separation of science, statistics, and engineering creates a problem that is present in all technology, and that while it seriously affects only a few branches of engineering at present, its impact upon all of engineering is both inevitable and imminent.

In the domain of reliability, the problem may be stated as follows: reliability engineering requires the use of the scientific method and (thus) statistics, but today's engineering graduate simply does not have sufficient training in these fields. The purpose of this paper is to show the intimate connection between science, statistics, and reliability engineering, and thereby to demonstrate the need for training engineers more thoroughly in the scientific method.

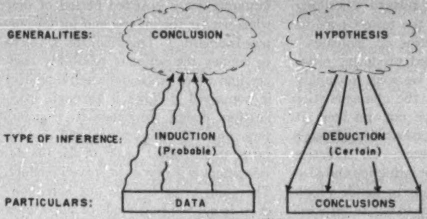
The Scientific Method

The purpose of science is the understanding of reality, and science has arisen historically from the human need to ad-

---

Paper 55-151, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing November 30, 1954.

J. C. Bear is with Aeronautical Radio, Inc., Washington, D. C.



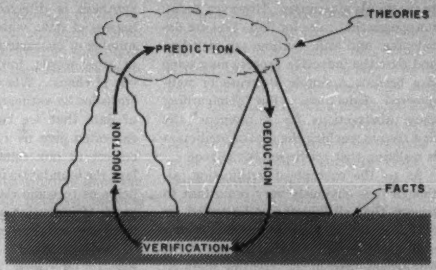


Fig. 1 (above). The two kinds of inference

Fig. 2 (right). The scientific method

just to this reality in order to survive and prosper. The scientific method is the process whereby we reach this understanding of reality, or, in short, it is the way science works, it is what scientists do. We must agree at the onset that this "reality" we speak of is not restricted to past events and that, in fact, the key to the problem of survival lies in our being able to predict the future with some measure of certainty. To help predict future events is the whole aim of science.[2] The only basis for a prediction is observation of past events; and since false or inaccurate predictions are harmful, predictions must be verified so that subsequent predictions will be more certain. From a preliminary standpoint, then, the three basic steps of the scientific method are: 1. Observation. 2. Prediction. 3. Verification.

Predictions have survival value only when expressed in terms of real events. Science is derived from facts and returns to facts for verification; also, since the human brain has but a finite number of brain cells in which it can retain information, it is imperative that all predictions be formulated in the simplest and most general terms possible. But before descending any farther into the details of the scientific method, we must first agree on the meanings of certain words.

The word "fact" is defined as "an event in reality." We shall assume reality to exist in a time-continuum so that events may be assumed predictable as well as observable. When we observe a fact we label it a "phenomenon"; and when we decide to use a phenomenon as a basis for formal thinking, we call it a "datum" or, in the plural ,"data."

The word "hypothesis" is defined as "a tentative proposition based upon facts." Notice that here we do not allow the word hypothesis to take on the meaning of an "a priori" or "self-evident" premise, for science concerns itself only with measurable reality. In other words, imaginary hypotheses are discounted.

The word "inference" shall be defined as "the logical process of forming a conclusion from given data or hypotheses." Inferences are therefore of two kinds: 1. inductive, where one starts with data, and 2. deductive, where one starts with hypotheses.[3] (It is more common to speak of deductions starting from premises than from hypotheses, but we prefer not to use the word "premise" because it carries the connotation of arbitrariness.)

Now consider the concept of "order" (as opposed to disorder or chaos). All our observations show us that no two events are exactly identical, i.e., that every "particular" event is unique. But we also find that several of these particular events may have some characteristics in common, so we group events into classes and then speak of the "general" characteristics of the whole class.

Using these definitions, we can now diagram the two thought-processes of the scientific method, induction and deduction (see Fig. 1). It may help to think of both types of inference as pyramids, with foundations imbedded in particulars and with summits in a cloud of generalities. Induction proceeds upward from the particular to the general, and since generalities are (like clouds) always somewhat hazy and approximate by nature, induction itself is an uncertain process that implicitly involves a statement of probability. Deductions, on the other hand, proceed from the general down to particulars, following established rules of logic. Deduction, therefore, is a "certain" process in the sense that the conclusions follow necessarily from the hypothesis (although not in the sense that the conclusions are certain to be true or certain to be verified).

The scientific method depends equally upon both of these processes of reasoning. It had its historical beginning in induction, which is to say, it sprang from the primitive, animal need for prediction. But induction suffered a terrible setback in the Middle Ages when man's conclusions were given to him as a priori religious principles and he was put to torture for deviation. During that era, man amused himself with deductive exercises (often reasoning from absurd premises, it may be added), so that the rules for deduction developed far in advance of those for induction.[4] In fact, it is still impossible today to delineate complete step-by-step techniques for proceeding from the particular to the general. Modern statistical theory is currently leading the way in the development of these procedures for inductive inference.[5]

It is fundamental that while science periodically soars off in inductive ascensions, it must always return to reality for verification. This process of verification is recognizable to the engineer as nothing more than a "factual feed-back loop." So now, having taken notice of the circularity of science, we can attempt a definition of its method: the scientific method is a self-correcting process of inquiry based upon fact and aimed at verifiable prediction. We can diagram this statement by tying together the two pyramids of induction and deduction as shown in Fig. 2. At the top, we put the "prediction" that is required to shape the conclusion of the inductive inference into a working hypothesis; and at the bottom, we use experimental "verification" to tie the deductive consequences of a hypothesis to the factual foundation of a new induction. It would be misleading, however, to give the impression that a scientific theory is evolved by a single induction and is established by a single deduction. Rather, these two modes of thinking accompany each other in the process; the diagram shows only the very general steps.[6] Deduction comes fairly naturally to us, while induction seems difficult; indeed, it is still quite common to think of scientific induction as being restricted

ol. 74
. 1
955

to the minds of geniuses. However, today it appears that rules for induction are developing and will continue to develop, and that the inductive process may some day become a simple staircase of well-directed deductions. But disregarding these abstractions for a moment, the first step in reaching the goal of prediction is realizing that prediction is the goal.

As to the concept of verification, although it is obviously very important to science that all hypotheses be tested by experimental measurement, it was not until quite recently in man's history that this became part of the scientific method. In 1591, Galileo climbed the Leaning Tower of Pisa and dropped a 10-pound weight and a 1-pound weight together from the top. When it was observed that they struck the ground simultaneously, we say that modern science began.[4] Galileo had returned to the facts for verification of his hypothesis and had thus closed the loop. This act served to warn all scientists not to propose unverified hypotheses, and thus rose the need for objectivity and skepticism, ingredients that have become fundamental to the pursuit of science. As Darwin said, "I have steadily endeavored to keep my mind free so as to give up any hypothesis, however much beloved (and I cannot resist forming one on every subject), as soon as the facts are shown to be opposed to it."[4]

The Role of Statistics

We have already defined statistics as the technology of the scientific method. While probability theory dates back to the 17th Century, statistics is a relatively new field and is still rapidly developing. For this reason, there is considerable disagreement among statisticians as to just what the function of statistics should be.[1,2,5] There seems to be agreement, however, that the mathematical "models" supplied by probability theory are derived deductively from postulates of greater generality, while the statistics that we try to fit to these models are derived inductively from data of less generality. We shall not argue the finer philosophic points here, but merely show where and how statistics fit into the scientific method as currently used in reliability engineering.

Consider again the process of induction (Fig. 1). Induction implicitly involves a statement of probability and thereby indicates (as mentioned before) that statistics is employed. Induction starts from data, and statistics supplies the techniques for marshalling these data. The whole mechanism of statistical synthesis is directed toward describing masses of data with the fewest possible number of characteristics, or "statistics," in other words, toward producing order out of chaos. When we wish to employ statistics to estimate the characteristics of data that we have not yet seen, it can often give us a probability of being correct in our estimates. So statistics lays the foundation for induction and also helps us to some extent to decide which paths of induction are likely to lead toward successful generalization.

Now, consider the deductive process in Fig. 1. Of course, all the practical applications of science fall into this area of reasoning from the hypothesis or theory down to particular conclusions. (Traditionally, this area includes the field of engineering.) We do not usually need the help of statistics to derive these specific applications, for the well-developed rules of deductive logic will suffice. But whenever we decide to test the validity of a hypothesis, we are obliged to call on statistics, for it is statistics that gives us the design of our critical experiment. This experiment must be very carefully planned so as to eliminate or randomize all the variables not under test. That the hypothesis may be tested in the most efficient and exact manner possible. Thus it is the process of verification that specifically requires statistical design; but since a hypothesis is vitally affected, we would be wise to consult the statistician as to the accuracy of our deductive conclusions as well.

The field of statistics, then, is primarily concerned with the foundations of inductive inference and with experimental verification, and it operates in the lower areas of the structure of pyramids (Fig. 2). These lower areas are concerned with facts or data, and it is in these regions that we should expect to find statistics; after all, the field of statistics has arisen to meet the need for the systematic handling of large bodies of data. In the layman's definition, a statistician "analyzes data and tells us what they mean." It must be emphasized that at present the statistician is limited by his theoretical development to the phases of the scientific method that deal with facts or data. Someone else must formulate the hypotheses; someone else must make the predictions. In reliability studies, this "someone else" is an engineer.

Reliability Engineering

The reliability of an electronic product may be defined as the "probability that the product will give satisfactory performance for a specified period of time when used in the manner and for the purpose intended."[6] This definition will serve to convey the scope of reliability engineering, which originates when a product is conceived, follows it through specification, design, development, manufacture, use, and maintenance, and then returns to the point of origin for the development of a new or improved product. The point to be brought out here, however, is not the magnitude of importance of the field, but that this definition is stated in the form of a prediction. The objective of reliability engineering is to present the design engineer with a systematic method for predicting the degree of satisfactory performance he may expect from his product. Of course, there are many such methods already in use in the form of equipment evaluation programs that freed failure information back to the designer. Certainly this concept is not new, and the problem of improving reliability is one of degree. But the degree is not small; considering the increasing reliance placed on electronic equipments, there appear to be several orders of magnitude of room for improvement. Specifically, we want to provide the designer with not only a method for improving existing products, but also a means for forecasting the reliability of new products and, further, with a measure of the certainty of his prediction.

The reliability engineer is faced with the formidable task of developing these long-range predictions. As has been indicated, his most efficient approach appears to lie in using the scientific method, which has evolved over the years for the exact purpose of developing valid predictions. The steps he should take, already outlined, are summarized as follows:

1. Observation—Gather data which will give a measure of the performance of the product and will relate this performance to the time of its operation, the conditions of its use, and the satisfaction of its users.
2. Prediction—Assemble these data and formulate generalizations in the form of brief and simple hypotheses that describe not only the data gathered but also data not yet gathered.
3. Verification—Reason out the logical consequences that these hypotheses suggest, then test them by controlled experiment.

The reliability engineer who follows these steps may be guaranteed success, at least on a logical basis, but the process is not as simple as that.

The Major Problem

First, the engineer has not, as a rule, been subjected to training in the scientific method or in the special field of statistics. He is therefore inclined to use indiscriminately words such as "significance," "average," "probable," "confidence," "frequency," "normal," "representative," etc. He quite justifiably feels frustrated when the statistician objects to his statement. On the one hand he is being paid to produce results, and on the other hand he is criticized for doing so.

Second, even when the engineer has had formal training in the techniques of science, it does not follow that he possesses the imagination required to climb out of a miscellaneous pile of data and up to a valid generalization. It is not so much that an engineer does not like to generalize as it is that he does not always realize that it is necessary to state his generalizations in such a form that they can be tested and experimentally verified. Then, even while we know that these things must be done, it is sometimes hard to say exactly how to proceed. We can only submit that the engineer can become adept in these matters only through a study of science, its method, and its language. Until the time when university curriculums and on-the-job training have met this need, the statistician-consultant is certain to remain the reliability engineer's closest ally.

Conclusion

This paper has attempted to show the intimate relationship between science, statistics, and reliability engineering, and to point out certain incompatibilities between the engineer's educational background and the complexity of the problems he must solve. It may appear that in discussing only the more theoretical aspects of the problem we have ignored the importance of the practical approach and of practical experience. Certainly this was not intended. Good engineers are those who have learned through experience to apply the scientific method, either knowingly or intuitively.

It is emphasized that reliability is an engineering problem: our predictions must be directed toward the design engineer who will use them, and they must be formulated in engineering terminology. Thus, while the specialist or consultant is most valuable and necessary, in the long run it is the engineer who must carry the load. In some instances, however, progress in the improvement of reliability may have to pause for a moment while the reliability engineer stops to think—about thinking.

References

1. INTRODUCTION TO THE THEORY OF STATISTICS (book), **A. McF. Mood.** McGraw-Hill Book Company, New York, N. Y., 1950, p. 1.

2. THE DESIGN AND ANALYSIS OF EXPERIMENTS (book), **O. Kempthorne.** John Wiley & Sons, Inc., New York, N. Y., 1952, p. 5.

3. AN INTRODUCTION TO LOGIC AND SCIENTIFIC METHOD (book), **M. R. Cohen, E. Nagel.** Harcourt, Brace and Company, New York, N. Y., 1934.

4. A HISTORY OF SCIENCE (book), **W. Dampier.** The MacMillan Company, New York, N. Y., 1936.

5. WHAT IS PROBABILITY **R. Carnap.** *Scientific American,* New York, N. Y., vol. 189, no. 3, Sept. 1953.

6. INVESTIGATION OF ELECTRON TUBE RELIABILITY IN MILITARY APPLICATIONS. *General Report No. 1,* Aeronautical Radio, Inc., Washington, D. C., Jan. 4, 1954, p. 1.

---

l, 74
1
55

A Supplement to the Brune Synthesis

F. M. REZA
NONMEMBER AIEE

ONE OF the most elegant contributions to electric network theory was made by Brune[1] in 1931. On the basis of work by earlier contributors, mainly Foster[2], W. Cauer,[3–5] and Fry.[6] Brune proved a fundamental theorem:

"The driving-point impedance of any network made of passive finite lumped linear elements (resistors, capacitors and inductors) is a positive real function of the complex frequency. Reciprocally, given any such function one could find a network configuration consisting of a finite number of linear passive elements having that function for its driving-point impedance."

For the synthesis of positive real functions (prf), Brune suggested the graph of a grounded ladder that correlates the synthesis problem to the expansion of a prf into a continued fraction. The imposition of a particular graph (which is rather a necessity) frequently caused Brune to have negative inductors in his synthesis procedure. However, he was able to prove that the negative inductors occur in a pattern that is equivalent to employing perfectly coupled coils.

Since, from a practical standpoint, the frequent use of perfectly coupled coils is not desirable, there has always been a great desire to eliminate this element. As a step toward this goal, Bode[7] suggests resistance padding. This simple observation has perhaps a deeper significance, since it may imply that proper usage of extra resistances in synthesis relieves the use of transformers. E. A. Guillemin has used in unpublished material some practical methods for eliminating the transformers in special cases. The first existence proof that a prf could be synthesized without transformers was given by Bott and Duffin,[8] whose paper stimulated considerable interest in network circles for further investigation of the problem. At present, it seems that the general problem of the replacement of the transformer in network synthesis with R, L, and C will be one of the most interesting topics of research in network theory.

The aim of this paper is to present a general discussion of the problem of the transformer in 1-terminal-pair networks. The main results have been outlined elsewhere.[9]

Outline of Positive Real Functions and Brune Synthesis Procedure

PRELIMINARY CONSIDERATIONS

The notion of impedance in electrical engineering is a familiar one. For instance, the impedances of a resistor, a capacitor, and an inductor are known to be respectively R, $(1/j\omega C)$, and $j\omega L$ for an applied voltage of frequency ω. Since actually dealing with complex quantities is involved, a complex variable $s=\sigma+j\omega$ may be conveniently defined as the complex frequency. The use of a complex variable s simplifies the study of the circuit and, when dealing with sinusoidal

---

Paper 55-140, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted September 10, 1954; made available for printing December 6, 1954.

F. M. REZA is with the Electrical Engineering Department and the Research Laboratory of Electronics of the Massachusetts Institute of Technology, Cambridge, Mass.

This work was supported in part by the Signal Corps; the Office of Scientific Research and Development Command; and the Office of Naval Research.

The author wishes to thank Professor W. R. LePage of Syracuse University for reading the manuscript, and Professors E. A. Guillemin, R. M. Foster and Dr. O. Brune for their continued interest in this work. The material for the Appendix has appeared in the Transactions of the Institute of Radio Engineers Professional Group on Circuit Theory.[11] See also reference 13.

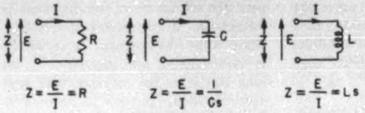


Fig. 1. Driving-point impedances of the three basic linear passive elements

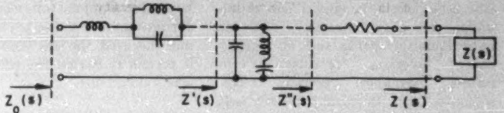


Fig. 3. Reactance, susceptance, and resistance reduction in the Brune synthesis

stimuli, we may finally replace s by $j\omega$. (See Fig. 1.)

Note that the three impedances R, $(1/Cs)$, and Ls are real functions, and their real parts are positive wherever the real part of s is positive

$$ReR = R \text{ essentially positive} \tag{1}$$

$$Re\frac{1}{Cs} = \frac{\sigma}{C(\sigma^2+\omega^2)} > 0 \quad \text{for } \sigma > 0 \tag{2}$$

$$ReLs = L\sigma > 0 \quad \text{for } \sigma > 0 \tag{3}$$

These properties of "realness" and "positiveness of the real part" are two basic properties that characterize all lumped linear passive networks. Indeed, the driving-point impedance $Z(s)$ of a random combination of these elements, as shown in Fig. 2, will preserve these properties. When a current of frequency ω is injected in the network of Fig. 2, the active power delivered to the network will be proportional to the $ReZ(j\omega)$, which must then be a nonnegative quantity.

POSITIVE REAL FUNCTIONS

A rational function $Z(s)$ is said to be a prf if

$Z(s)$ is real for real values of s
$ReZ(s) \geqslant 0 \quad \text{for } Res \geqslant 0$

Theorem 1: If $Z(s)$ is a prf, then $Z(s)$ has no poles in the right half of the s-plane. Poles on the boundary, if any, must be simple and with positive residues.

Since if $Z(s)$ is positive real and consequently $[Z(s)]^{-1}$ is also pr, it can be seen that in theorem 1 the term pole of $Z(s)$ can be interchanged with pole of $[Z(s)]^{-1}$.

Theorem 2: If $Z(s)$ is a prf then the minimum of the harmonic function $ReZ(s)$ in the right half-plane occurs on the axis of imaginaries.

The Schwarz lemma: If $W(\lambda)$ is a function regular in the unit circle $|\lambda| < 1$, and if $|W| < 1$ for $|\lambda| < 1$, and finally $W(0) = 0$, then either

$$|W| < |\lambda| \text{ in the unit circle} \tag{4}$$

or $W(s)$ must be a function of the form $W(s) = Cs$ with $|C| = 1$. Using this fundamental lemma of Schwarz in a polar system, Brune established the following interesting theorem.

Theorem 3: If $Z(s)$ is a prf

for $0 \leqslant |\text{Arg } s| \leqslant \frac{\pi}{2}$

then

$$|\text{Arg } Z(s)| \leqslant |\text{Arg } s| \tag{5}$$

The content of this theorem may be better appreciated in the following form

If $Z(s) = Ls$, where L is a positive constant, then $\text{Arg } Z(s) \equiv \text{Arg } s$. If $Z(s)$ is any other prf, then

$$|\text{Arg } Z(s)| < |\text{Arg } s| \quad \text{for } 0 < |\text{Arg } s| < \pi/2 \tag{6}$$

$$\text{Arg } Z(s) \equiv \text{Arg } s \quad \text{for Arg } s = 0 \tag{7}$$

$$|\text{Arg } Z(s)| \leqslant |\text{Arg } s| \quad \text{for } |\text{Arg } s| = \pi/2 \tag{8}$$

Equations 6, 7, and 8 show that if $Z(s)$ is not of the form $Z(s) = Ls$, then $\text{Arg}|Z(s)| - |\text{Arg } Ls|$ cannot become zero within the right half-plane unless perhaps on the real or imaginary axis. In other words, the function $Z(s) - Ls$ cannot have any zeros within the right half-plane except for the positive real and imaginary axis.

Theorem 4: If $Z(s)$ is a prf and L a positive constant, then $Z(s) - Ls$ (which is not necessarily a positive function) cannot have more than 1 zero within the right half-plane (on the positive real axis).

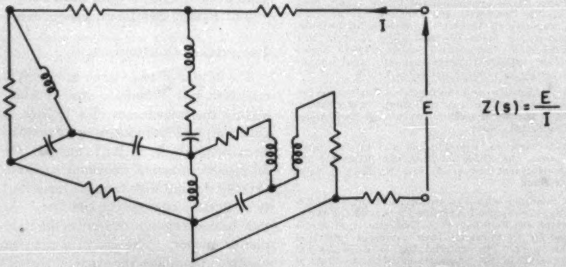


Fig. 2. Driving-point impedance of a general RLC network with mutual coupling

These four theorems, which outline the basic properties of the prf can be found in reference 1 or in texts related to network theory, e.g., reference 10, chapter 6, or reference 4.

BRUNE SYNTHESIS PROCEDURE

Step 1—Reactance Reduction

In conformity with theorem 1, if all j-axis poles of a given prf $Z_0(s)$ are properly removed, the remainder function $Z'(s)$ will be pr. (See Fig. 3.)

$$Z_0(s) = L_\infty s + \sum \frac{A_k s}{s^2+\omega_k^2} + Z'(s) \tag{9}$$

Step 2—Susceptance Reduction

Reconsider step 1 for the function $[Z'(s)]^{-1}$

$$[Z'(s)]^{-1} = C_0 s + \sum \frac{A_k' s}{s^2+\omega_k'^2} + [Z''(s)]^{-1} \tag{10}$$

Repeat steps 1 and 2 until no poles or zeros of the final remainder function $Z''(s)$ lie on the axis of the imaginaries.

Step 3—Resistance Reduction

Study the $Re[Z'''(j\omega)]$ and remove its minimum R in conformity with theorem 2.

$$Z(s) = Z'''(s) - R \tag{11}$$

It is possible that the synthesis of certain prf does not require any other steps. In such cases no mutual couplings are required in the Brune synthesis.

Step 4—Brune Cycle

When confronted with a $Z(s)$ such that the minimum of the $Re[Z(j\omega)]$ is zero and $Z(j\omega_0) = jL_1\omega_0$, remove a series inductance L_1; if L_1 is positive, then $Z(s) - L_1 s$, by virtue of theorem 4, ceases to remain a positive function and will have a positive real zero. Nevertheless, the nonpositive function $Z(s) - L_1 s$, which has a pair of zeros on the axis of imaginaries, suggests the application of step 2, followed by step 1. This final step introduces a negative inductor L_3. Brune proves that even with $L_3 < 0$, L_1, L_2, and L_3 collectively define a perfect transformer since

$$L_1L_2 + L_2L_3 + L_3L_1 = 0 \tag{12}$$

This step is called the Brune cycle (Fig.

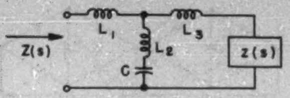


Fig. 4. Brune cycle

4). If $L_1<0$, $Z-L_1s$ will be pr, and steps similar to those given in the foregoing will lead to a Brune cycle, where $L_1<0$, $L_2>0$, $L_3>0$, and equation 12 is also valid; hence a physical realization in the form of a perfect transformer is immediate.

The remainder function $z(s)$ will be synthesized in the same manner. References 1 and 13 present elaborate proof of the positive real character of $z(s)$.

Synthesis Without Ideal Transformers

In the Brune synthesis of a driving-point impedance $Z(s)$, when $ReZ(j\omega)$ has a finite minimum along the axis of imaginaries (real frequencies), one is confronted with closely coupled coils. Let $Z(s)$ be such a function, whose minimum real part has been removed. As a result of the Brune cycle it is known that the removal of elements L_1, L_2, L_3, and c, shown in Fig. 4, will lead to a positive real function $z(s)$ which may or may not contain perfect couplings in its Brune configuration. From Fig. 4 is derived

$$\frac{1}{Z-L_1s}=\frac{cs}{cL_2s^2+1}+\frac{1}{L_3s+z} \tag{13}$$

$$Z_1(s)=Z-L_1s=\frac{(L_3s+z)(cL_2s^2+1)}{[c(L_3+L_2)s^2+1]+csz} \tag{14}$$

From theorem 4 we know that $Z-L_1s$ may have only 1 real zero in the right half-plane. Denoting this zero by k

$$Z(k)-L_1k=0 \tag{15}$$

and from equation 14

$$z(k)+L_3k=0 \tag{16}$$

In an effort to avoid the obstacle of the perfect coupling, it seems natural at first to attempt the Brune synthesis of the reciprocal function on the admittance basis (reference 11, p. 368).

$$Y_2=\frac{1}{Z(s)}-As$$

where A is a positive constant. Since $[Z(s)]^{-1}$ is a prf with $Re[Z(j\omega)]^{-1}=0$ then, from theorem 4, it can be seen that Y_2 has all of its poles and zeros in the left half-plane or the boundary, except for 1 positive real zero (since L_1 is assumed to be positive). A may be selected such that the positive real zeros of Z_1 and Y_2 coincide. Then

$$\frac{1}{Z(k)}-Ak=0,\ A=\frac{1}{L_1k^2} \tag{17}$$

Now, for this value of A, the ratio Z_1/Y_2 will have no poles or zeros within the right half-plane.

$$\frac{Z_1}{Y_2}=\frac{Z-L_1s}{1-AsZ}Z=\frac{Z-L_1s}{L_1k^2-sZ}L_1k^2Z \tag{18}$$

It can be shown that the function

$$Z_3(s)=k\frac{Z-L_1s}{k^2L_1-sZ},\ L_1>0 \tag{19}$$

is a prf. In fact, if $Z(j\omega)=U+jV$, then

$$ReZ_3(j\omega)=k\frac{L_1U(k^2+\omega^2)}{(k^2L_1+\omega V)^2+\omega^2U^2},\ L_1>0 \tag{20}$$

which is evidently positive, for all real finite values of ω. The network configuration of Fig. 5 is an immediate consequence of equations 19 and 20.

$$Z=kL_1\frac{kZ_3+s}{k_1+sZ_3}=\frac{1}{\dfrac{Z_3}{L_1k}+\dfrac{1}{L_1s}}+\frac{1}{\dfrac{s}{L_1k^2}+\dfrac{1}{kL_1Z_3}} \tag{21}$$

Note that Z_3 is not only a realizable function but is also simpler than the original function $Z(s)$. This is because Z_3 has a pair of zeros at $s=\pm j\omega$; hence its synthesis following Brune procedure allows the susceptance reduction (step 2). The continuation of Brune synthesis of function $Z_3(s)$ may be again confronted with perfect transformers. In such case the cyclic procedure described in the foregoing may be continued. This result was first introduced in reference 9.

Conversion Cycle

The next task is to determine the element values of Fig. 5. Since $Z_3(s)$ has a pair of zeros on the axis of imaginaries, step 2 of the Brune synthesis suggests that according to the partial fraction expansion of equation 10, the corresponding residues for these poles of $1/Z_3$ be removed.

$$\text{Residue of }\left(\frac{1}{Z_3}\right)\text{ at }\omega=\frac{\pm j}{(L_2/c)^{1/2}}=\frac{L_1(k^2L_2c+1)}{2kL_2^2c} \tag{22}$$

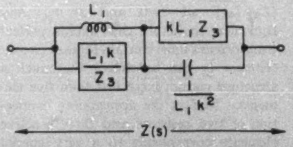


Fig. 5. Equivalent structure for a Brune cycle

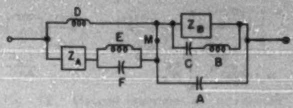


Fig. 6. Balanced bridge conversion cycle for a Brune cycle

Let the corresponding elements be an inductor and a capacitor with respective values L_v and c_v. Then

$$L_v=\frac{L_1(k^2L_2c+1)}{kL_2}$$
$$c_v=\frac{kL_2^2c}{L_1(k^2L_2c+1)} \tag{23}$$

(The reader is assumed to be familiar with the so-called Foster expansions of reactances. See reference 11.) This suggests the equivalent configuration of Fig. 6 for Fig. 5. Equations 23 and 21, along with the configuration of Fig. 6, show that

$$Z_B=\frac{L_1k}{Z_3}-\frac{L_1^2(cL_2k^2+1)s}{L_2(cL_2s^2+1)} \tag{24}$$

$$\frac{1}{Z_A}=\frac{1}{kL_1Z_3}-\frac{(k^2L_2c+1)s}{k^2L_2(cL_2s^2+1)} \tag{25}$$

or, after some algebraic operations, it may be found that

$$Z_B=\left(\frac{-L_1^2}{L_3}\right)\left(\frac{-L_3k^2-sz}{L_3s+z}\right) \tag{26}$$

$$Z_A=\left(-k^2L_3\right)\left(\frac{L_3s+z}{-sz-L_3k^2}\right) \tag{27}$$

Since $L_1>0$, $L_3<0$ has been considered, it can be seen from equations 19 and 20 and theorem 4 that Z_A and Z_B are prf, in spite of some negative values appearing in equations 26 and 27. Note that according to equation 16 the number k is the common zero of $L_3s+z(s)$ and L_3k^2+sz. If in the Brune cycle $L_3>0$, $L_1<0$, similar treatment of the problem on the admittance basis can be given.

From this discussion it is concluded that the structure of Fig. 6 should be considered a conversion of the Brune cycle. Indeed, each element A, B, C, D, E, F, Z_A, and Z_B of Fig. 6 is an explicit function of the elements appearing in the Brune cycle of Fig. 4. Hence the following steps are suggested as supplements to the Brune procedure.

SUPPLEMENT TO BRUNE PROCEDURE

Step 5—Element Values

When the first Brune cycle with ideal transformer is determined, then draw the conversion cycle (Fig. 6) and find its element values from equation 28.

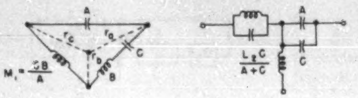


Fig. 7. A Δ-Y transformation applied to the balanced bridge conversion cycle

$$A=\frac{1}{L_1k^2} \qquad D=L_1$$

$$B=\frac{k^2L_2^2c}{k^2L_2c+1} \qquad E=\frac{L_1^2(k^2L_2c+1)}{L_2}$$

$$C=\frac{k^2L_2c+1}{k^2L_2} \qquad F=\frac{L_2^2c}{L_1^2(k^2L_2c+1)} \tag{28}$$

$$Z_A=k^2L_3\frac{L_3s+z}{sz+L_3k^2} \qquad Z_B=\frac{L_1^2}{L_3}\frac{L_3k^2+sz}{L_3s+z}$$

(For $L_1 < 0$, proceed on the admittance basis.)

Step 6—Continuation

Continue with the Brune synthesis of Z_A or Z_B. (Z_A is the reciprocal of Z_B within a constant multiplier.)

A Bridge Conversion Cycle

In this section will be described a conversion cycle with fewer elements than the cycle described in the preceding section. Two methods of solution will be discussed, one as an improvement of the previous cycle, the other as a rather independent solution that may have other possibilities.

METHOD A

The conversion cycle described in the previous section possesses the structure of a balanced bridge

$$L_1s\times\frac{L_1k^2}{s}=kL_1Z_3\times\frac{kL_1}{Z_3}=k^2L_1^2 \tag{29}$$

Hence the opening of the node M in Fig. 6, or the insertion of any passive impedance in that branch, will not affect the balance of the bridge; and inserting an inductor may be considered, with inductance $M_1=CB/A$ of such a form that the four elements (B, C, M, and A) also form a balanced bridge. Now we may apply a Δ-Y transformation as shown in Fig. 7. After some algebraic manipulations, the following results are obtained

$$r_a=\frac{1}{(A+C)s}$$

$$r_b=\frac{L_2cs}{A+C} \tag{30}$$

$$r_c=\frac{C\times L_2cs}{A(A+C)(L_2cs^2+1)}$$

Now r_c and the parallel resonant circuit of E and F could be combined to give a new parallel resonant circuit.

$$r_c+\frac{Es}{EFs^2+1}=\frac{s}{L_2cs^2+1}\left(\frac{CL_2c}{A(A+C)}+E\right) \tag{31}$$

This leads to the unbalanced bridge conversion cycle of Fig. 8 and equation 32.

$$D=L_1s$$

$$M=\frac{L_2L_3ck^2}{L_3ck^2-1} \qquad N=\frac{L_2c}{M}=c-\frac{1}{k^2L_3}$$

$$E_1=\frac{L_1^2(1+cL_3k^2)^2}{L_2(1-cL_3k^2)} \quad F_1=\frac{L_2^2c(1-cL_3k^2)}{L_1^2(1+cL_2k^2)^2} \tag{32}$$

$$Z_A=\frac{k^2L_3(L_3s+z)}{sz+L_3k^2} \quad Z_B=\frac{L_1^2k^2}{Z_A}$$

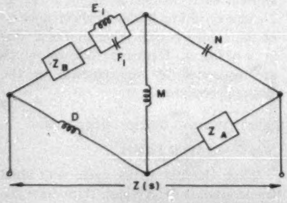


Fig. 8. An unbalanced bridge conversion cycle

Note that all elements appearing in Fig. 8 are physically realizable without the use of perfect transformers.

METHOD B

The search for a conversion cycle may be begun by considering a biquadratic driving-point impedance function presenting one Brune cycle such as that shown in Fig. 4, with $z(s)=r$. Then

$$Z(s)=\frac{[(L_1+L_2)cs^2+1]r+(L_1+L_3)s}{[(L_2+L_3)cs^2+1]+rcs} \tag{33}$$

The Brune cycle representing equation 33 uses one resistor, three positive reactive elements, and one negative inductor. The existence of the perfect coupling ($L_1L_2+L_2L_3+L_3L_1=0$) allows the use of only one resistor r in a network which must behave, respectively, as r and $[(L_1+L_3)/(L_2+L_3)]r$ at $s=0$ and $s=\infty$. Any equivalent structure that avoids perfect transformers must contain at least two resistors. It can be shown that such a structure cannot have less than five elements. In fact the appropriate connection of two resistors and three reactive elements corresponds to a very special subclass of the general biquadratic $Z(s)$.

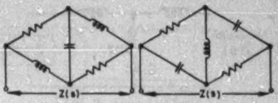


Fig. 9. A particular bridge for presenting a subclass of biquadratic positive real functions

Fig. 9 shows such configurations, but they do not cover the general problem. But there was in the preceding section a conversion cycle for the general biquadratic $Z(s)$ containing eight elements. This suggests a search for a structure containing two or three resistors and four or five positive reactive elements. The detailed solution to such a problem is not given here although it seems worthy of discussion. When some familiarity with the problem is established, it is rather natural to consider a structure similar to that shown in Fig. 10. In fact, if $Z(j\omega_0)=jL_1\omega_0$, then at $s=j\omega_0$ the two resistors must be deprived of dissipation. In the light of the next section two resonant and antiresonant tanks are first considered. The structure of Fig. 10 suggests itself as a possible solution.

The values of $Z(0)$ and $Z(\infty)$ determine the magnitude of the two resistances. Also, it is evident that for Fig. 10 to be the equivalent of Fig. 4 (for $z=r$) it must necessarily be assumed that

$$E_1F_1=L_2c \quad M_1N=L_2c \quad D=L_1 \tag{34}$$

The driving-point impedance pertaining to the bridge of Fig. 10 is a rational function $Z(s)=[N(s)/D(s)]$, where $D(s)$ and $N(s)$ are, in general, 5th-degree polynomials.

$$\begin{aligned}N(s)=(L_2c)^2aL_1rs^5&+[(M_1+E_1)L_1L_2c+\\ ar^2L_2c(L_2c+L_1N)]s^4&+[(M_1+E_1+\\ L_1)L_2cr+2L_1L_2car&+ENrL_1]s^3+\\ [(M_1+E_1)L_1+ar^2L_2car^2L_1N]s^2&+\\ [(M_1+E_1+L_1)r+arL_1]s&+ar^2\end{aligned} \tag{35}$$

$$\begin{aligned}D(s)=L_1(L_2c)^2s^5&+[(L_2c)^2r(a+1)+\\ L_1L_2crN]s^4&+[(M_1+E_1)L_2c+\\ 2L_1L_2c+ar^2L_2cN]s^3&+[L_2cr(2a+\\ 1)+rN(E_1+L_1)]s^2&+[(E_1+L_1+\\ M_1)+ar^2N]s&+ar\end{aligned} \tag{36}$$

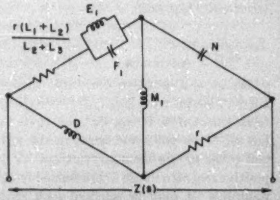


Fig. 10. Bridge conversion cycle for biquadratic case

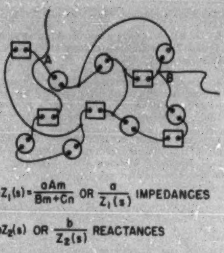

A
B
$aZ_1(s)=\frac{aAm}{Bm+Cn}$ OR $\frac{a}{Z_1(s)}$ IMPEDANCES
$bZ_2(s)$ OR $\frac{b}{Z_2(s)}$ REACTANCES

Fig. 11. General configuration made by random combination of two particular classes of prf's generating mutual coupling in the equivalent Brune configuration

where

$$a=\frac{L_1+L_3}{L_2+L_3} \tag{37}$$

Conditions for the identity of equation 29 with $[N(s)/D(s)]$ could be obtained in a straightforward but rather tedious computation. These conditions are found to be

$$D=L_1,\ M_1=c-\frac{L_3}{r^2},\ E_1=\frac{L_1^2(1+cL_3k^2)^2}{L_2(1-cL_3k^2)} \tag{38}$$

These considerations provide a bridge conversion cycle for converting a Brune cycle with perfect couplings into a cycle without couplings and containing fewer elements than the previous conversion cycle. To justify this general statement we can replace the resistor r in Fig. 4 by the function $z(s)$, which is the remainder of the Brune cycle. In Fig. 8, r and ar will be replaced, respectively, by Z_A and Z_B as found in the preceding section. It is a matter of algebra to check the fact that the identity of the two configurations is not altered by these operations. Thus one obtains results identical to that of method A. (The writer has established this identity algebraically but will omit that process for the sake of briefness.)

As an alternative to step 5, the following is offered as step 5(A): When the first Brune cycle with perfect couplings is determined, draw a bridge conversion cycle (Fig. 8) and find its elements from equation 32. (For $L_1<0$, proceed on the admittance basis.)

Concluding Remarks

The present trend of thought in the theory of communication engineering indicates that the different aspects of the problem of the synthesis of electrical networks will for some time be an important topic of theoretical and practical significance. A great variety of problems in other fields could be translated in terms of combination of linear network elements; analytical experience gained in the synthesis of networks is often useful in the design of other systems.

The problem of network synthesis is actually the problem of expanding a function into linear components with certain restrictions. Among the basic expansion methods, the expansion in continued fractions has an outstanding position. Since the Brune synthesis of a driving-point impedance consists of its continued fraction expansion, there is no doubt that its significance is not casual.

It is hoped that the present work will be of interest to those using the Brune synthesis, inasmuch as it offers a unified synthesis procedure using positive elements only. There are many other important related questions that remain unsolved. For instance, it was mentioned that if $Z(s)$ is a biquadratic prf, the Brune configuration of which presents a perfect coupling, an equivalent structure without couplings could be found with not more than seven elements. However, when $Z(s)$ requires n perfect couplings for its Brune configuration, the minimum number of elements for the equivalent structure without perfect couplings is not yet known.

Appendix. Generation of Driving-Point Impedance Functions Leading to Mutual Coupling[12, 13]

In the strictest interpretation of the Brune synthesis, the poles and zeros of the impedance function on the imaginary axis are first removed until the remainder function no longer has any more of them. Next, the smallest value of the real part of this impedance function for real frequencies must be removed. If the numerator of the $Re[Z(j\omega)]$ of the remainder has one or more roots at a finite nonzero real frequency, then the Brune synthesis requires mutual coupling. In such cases, the numerator of the $Re[Z(j\omega)]$, as a function of ω^2, will have one or more positive real roots of even multiplicity. In this study, "network with mutual coupling" refers to a driving-point impedance function having the last-mentioned property.

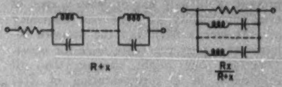


Fig. 12. General structure of a network block which, when embedded in a network, will generate mutual coupling in the equivalent Brune network

To identify the general class of positive real functions with this property, consider first the special class of driving-point impedances

$$Z_1(s)=\frac{Am}{Bm+Cn}=\frac{1}{1/R+1/[F(s)]} \tag{39}$$

where $m(s)$ and $n(s)$ are respectively even and odd parts of any Hurwitz polynomial of a degree higher than 1, and A, B, and C are positive real constants. $Z_1(s)$ can be synthesized immediately in the form of a resistance $R=A/B$ in parallel with a reactance $F(s)=Am/Cn$. Now consider the arbitrary reactance function $Z_2(s)$ and its inverse function $1/[Z_2(s)]$ combined with $Z_1(s)$ and $1/[Z_1(s)]$ in a random manner, as illustrated in Fig. 11.

$$Z(s)=\psi(Z_1, Z_2) \tag{40}$$

It is now possible to prove that the combination of these two classes of functions generates a network with mutual couplings. In fact, if $Z(s)$ is the driving-point impedance between any two nontrivial nodes A and B it has only to be proved that $Re[Z(j\omega)]$ has some positive ω^2 zeros of even multiplicity.

The $Re[Z_1(j\omega)]$ vanishes at the real frequencies $\omega_1, \omega_2, \ldots, \omega_n$ which are roots of $m(s)=0$. Assume a voltage of frequency ω_1 applied to the terminals A and B. Since $Z_2(s)$ is lossless, and the cubic blocks of Fig. 11 at such a frequency are either short-circuited or open-circuited, there must be no power dissipation. This means that ω_1^2 is a root of even multiplicity for $Re[Z(j\omega)]$. The original characteristic of the network $Z_1(s)$ is thus preserved through its block combination with lossless networks.

If $Z(s)=(m_1+n_1)/(m_2+n_2)$ is a prf representing mutual coupling in the Brune

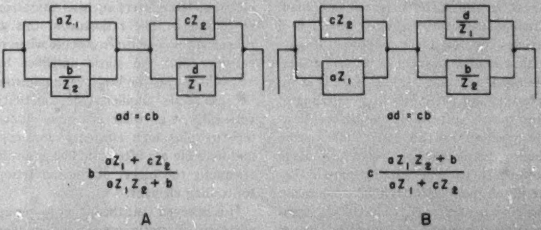


Fig. 13. A basic structure that presents mutual coupling in its Brune equivalent presentation

process, $Re[Z(j\omega)]$ has at least a root of even multiplicity at some real frequency ω_1. At such a frequency the network behaves as a pure reactance m_1/n_2 or n_1/m_2. Therefore, one might suppose that all resistances in a representation of $Z(s)$ which avoids mutual couplings would appear either in series with a parallel resonant circuit of frequency ω_1, or bridged across a series-resonant circuit of the same frequency. Thus, in general, it might be expected that these resistances would occur as a part of network blocks of special form, as shown in Fig. 12.

$$Z_1(s)=R+x \text{ or } Z_1(s)=\frac{Rx}{R+x} \tag{41}$$

The size of R and the impedance levels may vary for different blocks but the configurations are of the same general structure.

It is interesting to note that the parallel combination of $Z_1(s)$ with a reactance, as well as the series combination of $1/[Z_1(s)]$ with reactances, is generally recognizable in the Brune process without mutual coupling. The removal of poles and zeros of $Z(s)$ on the imaginary axis reveals such configurations except in particular cases of such mixed combinations as are shown in Fig. 13. These are among the general structures which cannot be broken down without mutual coupling in the Brune process. Without loss of generality let $Z(s)=s$; thus in Fig. 13 is recognized the structure used in the Bott and Duffin procedure.[8]

In the discussion of the converse property it has been assumed that the network is composed essentially of a combination of two kinds of functions: $Z_1(s)$ and reactances. For the more general case the possibility is readily conceivable of having other networks immediately in series or parallel with $Z_1(s)$. The general pattern of the somewhat "localized resistances" will still be retained.

For obvious reasons the function $Z_1(s)=1/[1+F(s)]$ might be called the generating function for the class of RLC impedances the Brune configuration of which represents mutual coupling.

It is important to note that in the synthesis of the impedance functions the freedom of choosing different graphs is possible. An impedance function requiring mutual coupling in a ladder structure may not necessarily require such coupling in a lattice structure.

In a more general attempt to find configurations avoiding mutual couplings, topological graphs are conceivable other than those of Figs. 12 and 13 for representation of driving-point impedance functions. In fact, if G is such a configuration, the resistances in G must be deprived of the flow of the current at applied voltages with those real frequencies which will make $Re[Z(j\omega)]$ minimum (except $\omega=0$ and $\omega=\infty$). Putting the case of Fig. 12 aside, other suitable graphs may be found by imposing the restrictions that at certain given applied frequencies, and only at those frequencies, the node voltage distribution in the graph must be such that any node pairs at terminals of a resistance will be equipotentials.

References

1. SYNTHESIS OF A FINITE TWO-TERMINAL NETWORK, **O. Brune.** *Journal of Mathematics and Physics*, Cambridge, Mass., vol. 10, 1931, p. 191.

2. A REACTANCE THEOREM, **R. M. Foster.** *Bell System Technical Journal*, New York, N. Y., vol. 3, 1924, p. 259.

3. DIE VERWIRKLICHUNG VON WECHSELSTROMWIDERSTAENDEN VORGECHRIEBENER FREQUENZABHAENGIGKEIT, **W. Cauer.** *Archiv Elektrotechnik*, Berlin, Germany, vol. 17, 1927, p. 355.

4. UEBER EINE KLASSE VON FUNKTIONEN DIE DIE STIELTJESSCHEN KETTENBRUECHE ALS SONDERFALL ENTHALT, **W. Cauer.** *Jahresbericht der Deutschen Mathematiker*, vol. 38, 1929, p. 63.

5. THEORIE DER LINEAREN WECHSELSTROMSCHALTUNGEN (book), **W. Cauer.** Becker and Erler, Leipzig, Germany, 1941.

6. THE USE OF CONTINUED FRACTIONS IN THE DESIGN OF ELECTRIC NETWORKS, **T. C. Fry.** *Bulletin*, American Mathematical Society, New York, N. Y., vol. 35, 1929, p. 463.

7. NETWORK ANALYSIS AND FEEDBACK AMPLIFIER DESIGN (book), **H. W. Bode.** D. Van Nostrand Company, Inc., New York, N. Y., 1945.

8. SYNTHESIS WITHOUT IDEAL TRANSFORMERS, **R. Bott, R. J. Duffin.** *Journal of Applied Physics*, New York, N. Y., vol. 20, 1949, p. 816.

9. SYNTHESIS WITHOUT IDEAL TRANSFORMERS, **F. M. Reza.** *Ibid.*, vol. 25, June 1954, p. 807.

10. THE MATHEMATICS OF CIRCUIT ANALYSIS (book), **E. A. Guillemin.** John Wiley & Sons, Inc., New York, N. Y., 1950.

11. A SUMMARY OF MODERN METHODS OF NETWORK SYNTHESIS (book), **E. A. Guillemin.** *Electronics*, New York, N. Y., vol. 3, ADVANCES IN ELECTRONICS, 1951.

12. QUARTERLY PROGRESS REPORT, **F. M. Reza.** Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., Jan. 15, 1953; April 15, 1954.

13. CONVERSION OF A BRUNE CYCLE WITH AN IDEAL TRANSFORMER INTO A CYCLE WITHOUT AN IDEAL TRANSFORMER, **F. M. Reza.** *Transactions*, Institute of Radio Engineers, New York, N. Y., vol. CT-1, no. 1, Mar. 1954, pp. 71–72.

Vo
Pt.
19

160,000-Ampere High-Speed Magnetic-Amplifier Design

ALLEN B. ROSENSTEIN
ASSOCIATE MEMBER AIEE

THE PAST 8 years, with the advent of improved semiconductor rectifiers and the availability of high quality magnetic materials have seen greatly renewed interest in the static magnetic amplifier. Current literature offers a number of articles on magnetic-amplifier theory, with a smaller number of reports available on design and application. It is interesting and perhaps a little puzzling to note that in spite of the fact that saturable reactors in the early 1900's were chiefly employed as relatively large power regulating devices, the present-day trend of magnetic-amplifier application has been toward small to moderate signal and power control. The controlled power design tendency has been to use low-level static magnetic amplifiers driving higher power rotating magnetic amplifiers such as rotatrols, amplydynes, and conventional generators or controlling electronic converters such as thyratrons or ignitions. The rationale of this arrangement is sometimes obscure when we consider that the static amplifier has been chosen for the low-level application because of its simplicity, dependability, versatility, etc., while the same characteristics plus high efficiency, low cost, and long life are often overlooked in determining the higher power converting-controlling element.

It is believed that the higher power application of static magnetic amplifiers may be limited, in part because of engineering concepts of magnetic-amplifier size, weight, power amplification, and speed of response that have carried over from the early saturable reactor days. It is not generally recognized, for example, that a static magnetic-amplifier regulated rectifier can inherently be made smaller and lighter than a comparable motor generator set or that the magnetic amplifier is capable of meeting and exceeding most of the performance characteristics of thyratrons and ignitions including that of half-cycle response. As a small effort to extend our static magnetic-amplifier power horizons, this paper will describe a high-power, fast-response magnetic amplifier which the author designed in 1950 and which has seen several years of highly successful service. An idea of the size of this magnetic amplifier may be

---

Paper 55-61, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 4, 1954; made available for printing December 9, 1954.

A. B. ROSENSTEIN is with the University of California, Los Angeles, Calif., and consultant to Inet, Los Angeles, Calif.

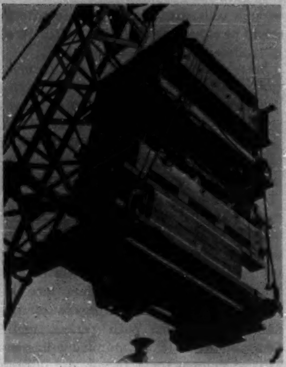


Fig. 1. Assembled unit hoisted for entanking

had from the fact that it contained over 20,000 pounds of copper, was capable of supplying up to 160,000 amperes (amp) of direct current and yet was controlled by 500 milliamperes (ma) of current. Figs. 1 and 2 give internal and assembled views of the amplifier. The over-all weight of the amplifier and accessories was 154,280 pounds. Exclusive of the transformer copper, 11,010 pounds of bus bars were used. The saturable reactor cores utilized over 4 tons of grain-oriented steel. The four steel containers each measured 11 feet high by 7 feet 4 inches by 5 feet 8 inches wide. Eight thousand gallons of oil were required to fill the system. The machine was the result of a complete survey of existing electric equipment to meet the specifications in Table I.

Equipment to meet the specifications were analyzed from the standpoint of efficiency, dependability, flexibility (it was desirable but not necessary to obtain higher voltages at reduced currents), and optimum approximation of the desired specification. Machinery considered included ignitrons, homopolar generators, d-c generators, and low-frequency alternators. The d-c magnetic amplifier was selected as being capable of most nearly approaching ideal performance.

General Considerations in Amplifier Design

Direct current was chosen to eliminate the high-current reactive drop of the load loop. However, inductive phase voltage drops are of prime concern in any low-voltage, high-current magnetic amplifier or rectifier. The effects of inductance will be discussed in detail in the section "Electrical Design." Here it can be seen that the inductive voltage loss renders a single unit construction impractical for 160,000 amp. This plus the need for transportability, and the additional flexibility, available from series-parallel connections, led to a decision to construct the amplifier in four equal sections, each rated 30,000 amp continuous. By changing the external bussing, four units allow adjustable voltage ranges of 8 to 56 volts, 4 to 28 volts and 2 to 14 volts. Fig. 3 gives the elementary block diagram for one unit.

Inductive voltage drops were further minimized through the use of the 3-phase, double half-wave circuit with interphase transformer. Each half-wave section then had to supply one half of the total current rating of the unit. The 3-phase, half-wave circuit offers maximum selenium cell economy and maximum over-all electrical efficiency for low-voltage operation. Double half-wave rectification gave improved transformer utilization, reduced telephone interference in the high-voltage supply, and allowed a conveniently balanced mechanical design.

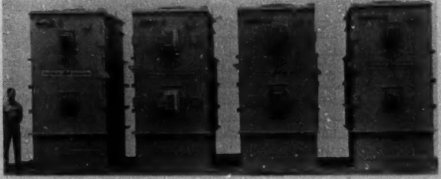
REACTORS

Table I. Specifications for Magnetic Amplifier

Input: 12,800 Volts, 3 Phases, 60 Cycles. Output Range: 2 to 14 Volts at 15,000 to 160,000 Amp. Control Adjustment Continuous from 2 to 14 Volts

Maximum Ratings	Volts	Amp	Output, Kw	Time Duration Seconds,
Continuous	14	120,000	1,680	
Intermittent	14	160,000	2,580	
Response time	1$^{1}/_{2}$ to 3 cycles			
Double-power cycle (to be repeated indefinitely)	7.5 10.6	119,000 168,000	890 1,780	9 1

Output control was achieved through the use of six self-saturating reactors per unit. These reactors (see Figs. 3 and 4) were located in the transformer secondary circuit with each reactor connected in series with one phase of the selenium rectifier to form a polyphase self-saturating d-c magnetic amplifier. The electromechanical characteristics of the chosen magnetic-amplifier circuit had to be balanced against those of the self-saturating magnetic amplifier connected in series with the transformer primary. For most high-current applications the mechanical problem attendant to feeding the heavy current through the reactors in the secondary circuit causes the primary doubler circuit to become mechanically attractive. In this particular case a unique electromechanical construction was evolved which caused the secondary reactor to be mechanically sound. The secondary magnetic amplifier introduces a voltage drop into the secondary circuit which necessitates a larger transformer than is required with a primary doubler. This transformer cost can be balanced against the additional selenium rectifiers required for the doubler circuit and the high-voltage insulation required in the power winding of the primary doubler. An



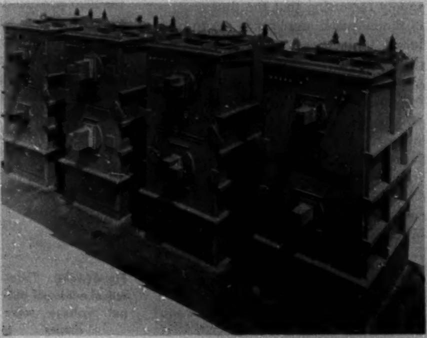


Fig. 2. Amplifier views from output side

additional important factor to be considered with the primary doubler is the d-c component of current which would flow in the transformer primary should there be any unbalance between the reactors of the doubler. The possibility of this direct current necessitates a slightly more liberal transformer steel allowance. The doubler also has a slower inherent speed of response. After carefully weighing the advantages of both reactor locations, it was decided that the integral mechanical construction offered by the secondary reactors overweighed any possible disadvantages.

Some of the mechanical problems of the reactors have been mentioned. Reactor space restrictions were also rather severe. Because of the extreme voltage range of the equipment, the reactors had to be large enough to absorb practically all of the transformer secondary voltage. Each reactor set had to be wound with copper capable of carrying up to 20,000 amp. The transformer secondary voltage and the inverse rating of the rectifier cells were directly dependent upon the degree of saturation that the control current could effect in the reactors. Therefore, residual reactance of the reactors had to be reducible at full ouput to a relatively small value.

The extremely fast response required for the double power cycle dictated that the magnetic-amplifier power gain be something less than 10,000/1. At the same time, the control of the double power cycle was expected to be very precise and completely dependable. By increasing the number of reactor control turns, the control current required by the magnetic amplifier could be reduced to a value that vacuum tubes could readily handle. With increased turns the voltages induced in the control windings increases the insulation problem and increases the required output voltage of the control supply (assuming that the time constant of the magnetic amplifier is not to be allowed to increase). For a given time constant, the reactor control power required becomes independent of the number of reactor control turns. Reactor control windings were selected to require a maximum of 500 ma allowing the use of moderate-sized transmitting tetrodes as control elements.

Although it was expected that a small amount of individual unit gain control would be necessary to final balancing of the units in parallel operation, strict quality control was exercised throughout the manufacture of the equipment to insure reactor balance within and between units over the complete operating range. Reactor cores were pregraded to within ±2½ per cent (%) before installation, and rechecked after installation.

Interphase

The interphase reactor design problem was more mechanical than electrical. Electrically, it was necessary to recognize that the shift in firing angle introduced by the magnetic gates (reactors) would cause a larger 180-cycle voltage to appear across the interphase than simple rectifier theory would predict. Mechanically, the problem was that of building a 30,000-amp interphase into a minimum of space while providing a convenient connection between the double-wye neutrals and the output terminal.

Selenium Cells

Although the selenium rectifier design presented some mechanical problems, they were largely based upon economic considerations. A low inverse voltage rating of the selenium cells would force extremely careful transformer-reactor design in order to allow maximum d-c output with a minimum transformer secondary voltage, but would permit higher efficiencies. High selenium cell inverse voltage ratings are usually accompanied by higher voltage drops in the forward direction and consequently reduced magnetic-amplifier efficiency. Since the magnetic amplifier was entirely self-contained and oil-immersed for outdoor operation, the heat produced by the losses had to be dissipated in a suitable heat exchanger. As sufficient cooling water was not available, oil-to-air heater exchangers were used. The size and cost of these exchangers was, of course, a direct function of the over-all amplifier efficiency. Although cells with higher inverse ratings might have been available, 33-volt inverse cells offered efficiencies of over 80%.

Electrical Design

As mentioned, the design of the magnetic amplifier revolved around the resolution of the rectifier inverse voltage rating, the required maximum d-c output voltage, and the total residual phase inductance of the system. The limitations upon the transformer secondary voltage are obvious. If 33-volt inverse selenium cells are chosen, the maximum allowable line-to-neutral secondary voltage will be $33/1.73 = 19$ volts, with sufficient load to cause the interphase to insure double half-wave action. With perfect rectifiers and reactors the minimum line-to-neutral voltage for a 14-volt d-c output must be $14 \times 0.855 = 12$ volts.

Design Procedure

The usual design procedure is to start with the minimum required a-c secondary voltage for the maximum desired d-c output (maximum voltage and current) and add the proper allowances to compensate for selenium cell drop, system *IR* voltage drop, and an assumed phase reactance potential loss. For the transformer voltage determined in the foregoing manner, the reactor size necessary to absorb sufficient voltage to allow minimum d-c output can be computed. With the reactor dimensions determined, its voltage drop at maximum load can be computed and checked against the original allowance used to determine the transformer secondary voltage. With this procedure a secondary transformer line-to-neutral voltage of 17.1 volts was established. The check of this value is as follows: An a-c forward voltage drop of 1.5 volts was assumed for the selenium cells at 40,000 amp per unit. The useful alternating voltage after rectifier drop becomes

$$17.1 - 1.5 = 15.6 \text{ volts rms} \qquad (1)$$

This voltage can now be converted to its equivalent d-c output value through division by the 3-phase, half-wave conversion factor

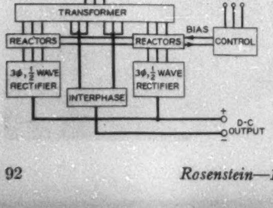
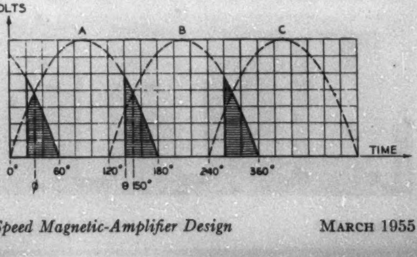


Fig. 3 (left). Elementary block diagram for one unit

Fig. 4 (right). Theoretical minimum output voltage wave shapes



Vol. 7
Pt. 1
1955

$$E_{do}=\frac{E_{ac}}{0.855}=\frac{15.6}{0.855}=18.2 \text{ volts} \qquad (2)$$

The total d-c IR drop of the unit including the equivalent transformer resistance was computed at 0.5 volt. Subtracting the IR drop and the desired maximum d-c output voltage from E_{do} will give ΔE_{do}, the total allowable commutating phase reactance voltage drop

$$\Delta E_{do}=18.2-0.5-14=3.7 \text{ volts} \qquad (3)$$

The general expression for ΔE_{do} is given by equation 4, where L includes all inductance per phase of each 3-phase, half-wave section, including the equivalent transformer leakage inductance, the reactor residual (i.e., saturated) inductance, and the total lead inductance per phase

$$\Delta E_{do}=\frac{p\omega LI_0}{2}=3fLI_0 \qquad (4)$$

where

f = supply frequency
L = total inductance
I_0 = average current (20,000 amp maximum per section)

Computing L for maximum load

$$L=\frac{\Delta E_{do}}{3fI_0}=\frac{3.7}{3\times 60\times 20{,}000}=1.03\times 10^{-6} \text{ henry} \qquad (5)$$

The magnitude of 1-microhenry inductance can be better appreciated by computing the inductance of a single-turn inductance with a radius of 1 meter and a conductor cross section of nominal area

$$L\cong\frac{\mu N^2\pi r}{2} \text{ for a simple circular coil}$$

$$L=\frac{12.56\times 10^{-7}\times 1^2\times \pi\times 1}{2}=2\times 10^{-6} \text{ henry} \qquad (6)$$

where

$\mu = 12.56\times 10^{-7}$ for free air
r = radius, meters

PHASE INDUCTANCE

Therefore, the total allowable phase inductance must be less than the inductance of a simple single-turn coil with a 1-meter radius. The components of the phase inductance were determined as follows: By careful design the equivalent reactance of the transformer was held to less than 1.0% or 0.05 microhenry. The lead inductance was estimated at 0.3 microhenry, although the complicated geometry made an exact determination difficult. The saturated reactor inductance was computed with approximately 400 ampere turns (a-t) of positive bias where the core curves of magnetic induction versus magnetic intensity (B-H) indicate that an additional 20,000 a-t will cause a core flux density change of 35,000 lines per square inch. The reactors were designed with one turn and a core area of 36 square inches. The reactor saturated inductance then became

$$L=\frac{N\Delta\phi}{I}=\frac{1\times 36\times 35{,}000\times 10^{-8}}{20{,}000}=0.63 \text{ microhenry} \qquad (7)$$

The total phase inductance = 0.05 + 0.3 + 0.63 = 0.98 microhenry or just within the allowable 1.03 microhenry. The steps taken to reduce the mechanical loop circumscribed by each phase and consequently the phase lead inductance to a minimum will be described in the section "Mechanical Design."

REACTOR DESIGN

The final electrical characteristics of the reactors were partially dictated by the combination of the low value of saturated reactor inductance required and the saturated characteristics of the reactor core material chosen. For a given kva, the saturated reactor inductance varies almost directly with the reactor turns. This means that a 2-turn reactor has one half the core area of the single-turn reactor, but possesses close to twice the saturated inductance, or in this case approximately 1.2 microhenrys. The final reactor dimensions can now be determined from consideration of the reactor requirements for minimum output voltage and current.

Minimum load conditions were 2 volts at 10,000 amp per unit. Final reactor design procedure calls for a check of the a-c reactor voltage that must be absorbed to achieve the desired minimum output against the maximum alternating voltage that can be absorbed by the chosen design. Since other voltage drops, including that of the selenium cell, still exist, the reactors do not have to absorb all the 17.1 volts available. The total alternating voltage that must be absorbed by all elements can be determined from a consideration of Fig. 4, which shows the output wave form of a perfect 3-phase, half-wave magnetic amplifier for 1 complete cycle. Since the average areas under the heavy wave give the d-c output, the firing angle of phase A can readily by computed for a 2-volt d-c output

$$\frac{3\times E_{max}}{2\pi}\int_{\theta}^{2\pi}\sin \times dx=E_{dc}$$

$$\int_{\theta}^{2\pi}\sin \times dx=\frac{2\times 2\pi}{3\times 1.414\times 17.1} \qquad (8)$$

$$-(\cos\theta-1)=17.3$$

$$\theta=180°-34.3°=145.7°$$

The volt radians absorbed by the reactor and associated elements of phase A must equal the area under voltage curve A less the shaded area when phase C or A have fired. For this particular machine, the IR and reactive drops will require phase C to fire well before time ϕ so the following figure for the volt radians that must be absorbed will be slightly conservative

$$\text{Area under wave } A=E_{max}\int_{\theta}^{2\pi}\sin \times dx=2\times 1.414\times 17.1=48.3 \text{ volt radians} \qquad (9)$$

$$\text{Area not absorbed}=2xE_{max}\int_{\theta}^{2\pi}\sin \times dx=2\times 1.414\times 17.1(1-34.30 \text{ degrees})=8.38 \text{ volt radians} \qquad (10)$$

$$\text{Area absorbed}=48.3-8.38=39.9 \text{ volt radians}$$

The equivalent rms voltage that must be absorbed is, of course, directly related to the volt radians.

$$\text{Rms voltage to be absorbed}=17.1\times\frac{39.9}{48.3}=14.1 \text{ volts} \qquad (11)$$

This voltage must be absorbed by the reactors, the inductive drop, the IR drop and selenium cell voltage loss.

With grain-oriented cores, the reactors are capable of sustaining a maximum flux density of 125,000 lines per square inch. With one turn and 36 square inches of core cross section, the reactors could each absorb

$$E=4.44Nf\phi\times 10^{-8}=4.44\times 1\times 60\times 36\times 128{,}000\times 10^{-8}=12 \text{ volts rms} \qquad (12)$$

At 5,000 amp per section, the lead and transformer induction drop becomes equivalent to 0.38 volt rms. The total selenium cell and IR voltage drop was estimated to be 1.3 volts. The total maximum voltage that can be absorbed adds up to 13.68 volt rms. It is to be noted that the maximum voltage of 13.68 volts which can be absorbed is slightly less than the 14.1 volts which must be absorbed to reduce the output to 2 volts direct current. Rather than increase the reactor size through the use of nonstandard cores, it was possible to use the two plus $2^1/_2$% transformer primary taps to reduce the secondary voltage to 16.3 volts rms. With the reduced secondary voltage it becomes necessary to absorb only 13.4 volts, leaving a small operating margin.

CONTROL DESIGN

The design of the reactor control windings requires correlation of the core material magnetization characteristics, core dimensions, control power source characteristics, desired power gain, desired response speed, and insulation requirements for the maximum inducted control-winding voltage. The reactor cores were composed of standard Hypersil *C* loops with an effective magnetic path of 33 inches. From the core magnetization curves, it was observed that 400 a-t corresponded to a point well beyond the knee of the saturation curve, while 100 a-t came well into the knee of the curve. The electronic control tubes available could conveniently supply 500 ma. A control winding of 800 turns with 500 ma would supply the desired maximum output, while 200 turns at 500 ma could supply the rated overload current at output voltages up to 11 or 12 volts. The minimum output could likewise be reached with 400 a-t of negative bias. To meet readily the full range of required operating conditions, the reactors were built with four 200-turn control windings with individual termination for series or parallel connection.

Since response was guaranteed to be within $1^1/_2$ to 3 cycles, the control system and control windings were carefully designed for maximum speed of response. Using the expression $L=N\Delta\phi/\Delta I$, the average inductance of the control when connected for 200 turns was computed to be 64 henrys. This figure dictated a 1,000-volt control power supply. The series resistance at 500 ma is 2,000 ohms and the control time constant becomes $64/2{,}000=0.032$ second or 1.92 cycles. Further to insure a maximum speed of response, the control current was regulated by series pentodes with a large amount of negative feedback. The pentodes act as constant current generators up to the point where rate of current increase approaches E_s/L. It should be noted that for operation at maximum output with 800 turns, the inductive control will not increase with the square of turns because of the nonlinearity of the magnetic core and the fact that at high outputs, one-third of the reactors are always conducting.

A quick check upon relative response for the conventional self-saturating magnetic amplifier can always be obtained from the over-all power gain of the amplifier. In this case an input change of 500 ma at 1,000 volts, 0.5 kw, was estimated to give the double power cycle with an 890 kw change in output power. A 1,980/1 power gain with less than 2-cycle response can be expected from a 3-phase self-saturating magnetic amplifier.[1]

In any large-power, fast-responding magnetic amplifiers, adequate instant high-voltage insulation must always be provided for the control windings. In this case, the maximum possible instantaneous voltage which could be coupled from the power winding to the control would be $2\times17.1\times200\times4=27{,}400$ volts per winding. Although 10,000-volt insulation was provided between control windings and to ground, the maximum possible coupled voltage was further reduced in several ways. First, each reactor was constructed of six Hypersil loops. The control winding encircled only two loops at a time so that each reactor had three sets of d-c coils of four 200-turn windings. For each 200-turn winding, the maximum coupled voltage becomes $2\times17.1/3\times200=2{,}280$ volts. Since the coupled control voltage of the reactors add phasorily when the winding of the various phases are connected in series, the net coupled voltage of all of the first 200 turns of the first section of all reactors would theoretically be zero. Since each reactor has three sets of windings, each 200 turns of control consists of $6\times3=18$ windings in series. To prevent an excessive inducted control winding potential from a rapidly reduced control current, Thyrite resistors were connected across the control input to each of the 30,000-amp sections.

To realize safely the full intermittent magnetic amplifier rating of 160,000 amp, it was necessary to balance carefully the magnetic characteristic of each reactor against the other reactors in its section. The magnetic characteristics of all sections were to be reasonably uniform so that less than 5% load unbalance between sections could be conveniently achieved. Each Hypersil loop was individually graded before and after installation. Characteristics were compared, and matched sets of six were selected. The final over-all reactor core characteristics were held to within $\pm2^1/_2\%$.

TRANSFORMER DESIGN

The low reactance and unusual mechanical requirements governed the transformer design. With a 12,800-volt input and 17.1-volt, high-current output, a thoroughly interleaved pancake construction was necessary. Because of the heavy currents, a single-turn secondary was used.

The necessity for relatively close coil spacing coupled with the coiling requirements of the rectifiers and further specification of a self-contained outdoor construction, made oil cooling and insulation attractive. With a closed cooling system, the possibility of high-voltage failures from the dirt and moisture that air cooling would accumulate on the transformer windings was eliminated. The transformer primary was tapped for $\pm2^1/_2$ and $\pm5\%$. The positive taps insured light-load 2-volt operation, while the negative taps were available for maximum output in case of rectifier aging.

RECTIFIER DESIGN

The rectifying elements of the magnetic amplifier were half-wave selenium stacks. Each 3-phase, half-wave section consisted of 6 half-wave stacks giving 2 stacks per phase and 12 stacks per unit. A total of 7,776 plates of 45 square inches (each $6^1/_4$ by $7^1/_4$ inches) were used to give an average oil-cooled current density of 0.344 amp per square inch.

Although 33-volt inverse cells were used, there was considerable speculation about the possibility of cell failure due to a high inverse voltage. It was believed that the cell manufacturer's inverse rating was conservative but required the failure of only 1 out of the 7,776 parallel plates to render the equipment inoperative. Further, if the equipment were energized with zero load, the inverse voltage would become twice the line-to-neutral voltage, or 34.2 volts instead of 1.73 times the secondary potential.

Four possible modes of cell failure, current overload, chemical, inverse voltage overheating, and inverse voltage barrier breakdown, presented themselves. Under continuous overload, the cell would overheat, the backing metal of the cell would melt, and the cell would be damaged or destroyed. However, with forced oil cooling, the possibility of individual cell overheating becomes negligible without materially increasing the oil temperature. Thermal overtemperature contacts were installed in the oil to turn the magnetic amplifier off well before a critical temperature could be reached.

Chemical elements in the oil could attach the selenium cells. A particularly pure neutral oil (without neutralizing chemicals) was specified and a regular check schedule established to insure that the neutralization number never exceeded 0.01 milligram of KOH. Above 0.01, the oil was to be discarded. Overheating caused by the high inverse currents of excessive inverse voltages would be prevented by the oil cooling.

Cell breakdown due to high inverse voltage failure of the barrier layer represented the only unknown factor in the cell reliability problem. To obtain some idea of the dependability of selenium cells under excessive inverse voltage, a random selection of twelve 33-volt cells was immersed in oil and subjected to slowly increasing inverse voltages. Other

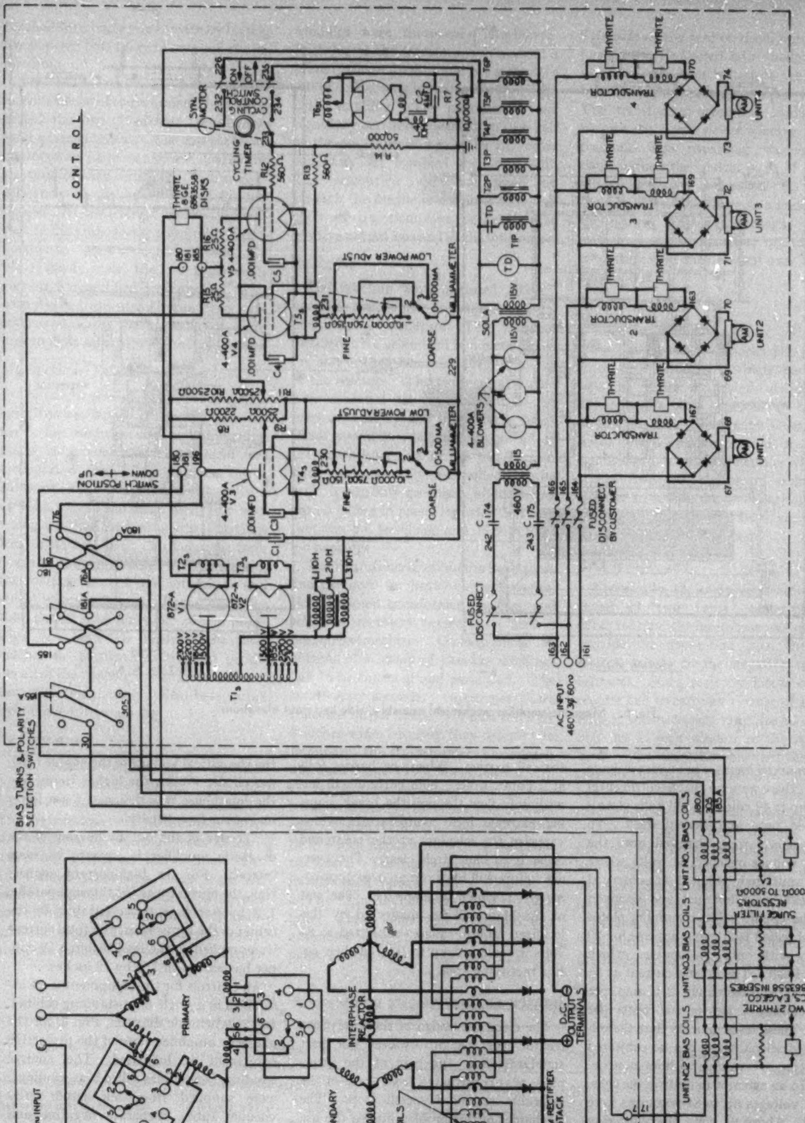

CONTROL
SYN. MOTOR
CYCLING CONTROL SWITCH
ON
OFF
CYCLING TIMER
THYRITE 8GE 6863558 DISKS
V5 4-400A
4-400A V4
4-400A V3
872-A V1
872-A V2
.001MFD
LOW POWER ADJUST
0-1000MA MILLIAMMETER
0-500 MA MILLIAMMETER
FINE
COARSE
10,000Ω 750Ω 150Ω
SWITCH POSITION DOWN UP
R1 R8 R9 R10 R11 R12 R13 R15 R16
2500Ω 2200Ω 560Ω
50,000
TD
T1P T2P T3P T4P T5P T6P
SOLA
115V
4-400A BLOWERS
460V
FUSED DISCONNECT
FUSED DISCONNECT BY CUSTOMER
AC INPUT 460V 3φ 60~
TRANSDUCTOR
THYRITE
UNIT 1
UNIT 2
UNIT 3
UNIT 4
L1 10H L2 10H L3 10H
2300V 2200V 2100V 1800V 1500V
BIAS TURNS & POLARITY SELECTION SWITCHES
UNIT NO. 1 12,800 VOLT, 3φ, 60~ INPUT
PRIMARY
SECONDARY
INTERPHASE REACTOR
D.C. OUTPUT TERMINALS
BIAS COILS
SELENIUM RECTIFIER STACK
UNIT NO. 2 BIAS COILS
UNIT NO. 3 BIAS COILS
UNIT NO. 4 BIAS COILS
TWO 2 THYRITE DISCS, E.A.G.E.CO. 6863558 IN SERIES
SURGE FILTER RESISTORS 2000Ω TO 5000Ω

Fig. 5. Magnetic-amplifier schematic with control, safety circuits omitted

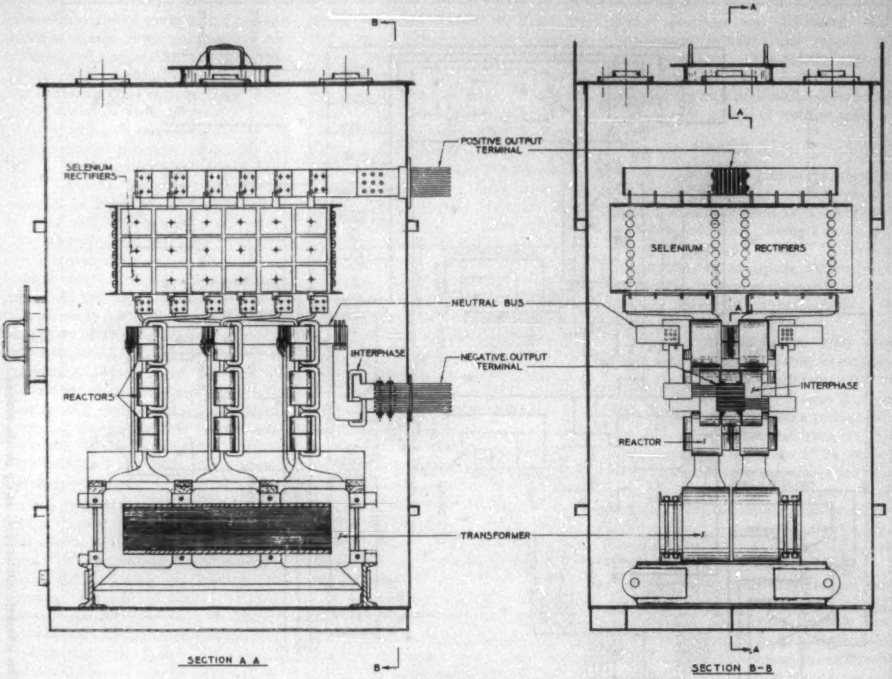


Fig. 6. Magnetic-amplifier mechanical assembly, side and front elevations

than a marked increase in inverse leakage current, there were no visible detrimental effects up to 42 volts. At 43 volts plate 9 sputtered once and cleared itself. No further effects were observed until the inverse voltage reached 51 volts where plate 2 sputtered. Plate 2 was observed to be extremely hot and was removed from the circuit. The remaining plates were subjected to 51 volts inverse for 20 minutes with only one sputter, which cleared itself. The inverse current at 51 volts was approximately 2.1 amp per plate, causing a loss of 107 watts per plate. These tests indicate that the oil-immersed selenium cell can be subjected to extreme overvoltage without destruction. In an attempt to destroy the cells, inverse voltages up to 84 volts rms were applied. Above 64 volts, the sputtering increased rapidly, but none of the cells were completely destroyed. Examination of the plates during and after the test indicates that very high increase voltages will cause random point failures in the cell barrier. When the barrier fails at a point, a very high current will immediately flow through the break, causing extreme local temperatures which vaporize the selenium at the point and cause it to blow itself clear. The average failure will blow out an area approximately 1/4 inch in diameter. The rest of the plate will be unaffected by this localized failure. Plates subjected to 84 volts inverse lost up to 50% of their active rectifying area.

Interphase Reactor

The electrical design of the interphase reactor was straightforward. For part standardization, the core of the interphase was constructed from three of the saturable reactor Hypersil loops. The mechanics of this problem forced the use of a 2-turn center-tapped construction. With the extreme phase shift introduced by the magnetic amplifier reactors, the maximum 180-cycle voltage across the reactor will be considerably greater than the theoretical voltage of the simple rectifier circuit. With the higher frequency, the interphase was designed for a lower maximum flux density.

The size of the output bussing of the magnetic amplifier is perhaps of some interest. For the high-current connection, the current was fed through parallel 1/2 by 6-inch water-cooled bus to the center of the array where the total current was supplied to the load through 12 copper bus bars each 1/2 by 12 inches.

All controls for the equipment were assembled in a single floor-standing cabinet. In the schematic diagram, Fig. 5, for the magnetic amplifier, none of the protective equipment is indicated. The control windings (sometimes called bias windings) were supplied from self-biased series vacuum tubes. Separate negative and positive bias tubes were provided. Current flow through the tubes was adjusted by varying the self-biasing cathode resistors. A motor-driven variable eccentric cam was provided to switch automatically the double or triple power cycle by biasing the positive bias tubes in and out of conduction.

The need for remote indication and the inconvenience of conventional shunts suggested the use of current transductors for d-c measurement. The current transductors were permanently installed within the tank of each section. The transductors enclosed the output bus of the section so that the average alternating current in the transductor was in direct proportion to the d-c output current. To use a linear scale, the a-c transductor current was rectified and applied to 5-amp d-c instruments. Thyrite disks were connected across the transductor windings to limit any inducted voltage surges.

MECHANICAL DESIGN

The mechanical design of the magnetic amplifier was largely predicated upon the needs for minimum-size and minimum-phase inductance. The equipment was transported from Los Angeles to the Midwest on special low-bed trucks. Even at that, road clearances along the way were such that the high-voltage bushings had to be removed. To keep the tank size and lead inductance to a minimum, the transformer, reactors, interphase, and stacks were made as an integral assembly that bolted into the tank as one piece. Figs. 6(A) and (B) illustrate the basic mechanical construction without detailing all of the bracing. Fig. 1 shows a section of the magnetic amplifier being lifted by a crane for entanking.

TRANSFORMER

The transformer, being the heaviest component, was used as the foundation for the magnetic amplifier assembly. The transformer was further laid on its side with the high-voltage leads and primary tap switch on the bottom. The low-voltage secondary leads were extended from the other side or what was now the top of the transformer. To maintain the desired low transformer reactance, the primary and secondary pancake windings were thoroughly interleaved. Each leg of the transformer contained a primary phase winding and two secondary sets of coils to supply the double-wye magnetic amplifier. The pancakes of one secondary set were collected on one half of the transformer leg, while the other secondary set occupied the other half. The separation was maintained up through the magnetic amplifier structure and the two sections were paralleled at the positive output bus and at the interphase.

REACTORS

The reactors were mounted directly upon the transformer and supported by a noninductive structure. The selenium rectifiers were assembled in a metal housing and in turn mounted directly on top of the reactors. The two neutral busses were mounted one on each side of the reactor assembly and fed into the interphase assembly which was supported on the low-voltage end of the transformer. After each assembly was bolted into its tank, it was only necessary, after putting the oil baffles in place, to install and connect the d-c terminals and high-voltage a-c bushings.

The limitations of minimum head room and the need for heavy-current connections between transformer, reactors, and selenium rectifiers presented a serious mechanical problem. On each phase the outputs of a group of pancake windings had to be collected and paralleled. The transformer secondary copper per phase collected into a bundle 3/4 inch thick by 9 inches wide. Rather than attempt to terminate the transformer secondaries and then connect to a heavy reactor bus, the transformer secondary windings were left approximately 40 inches long on each end. These ends were collected into 9- by 3/4-inch bundles to form vertical copper risers. The six Hypersil loops that formed each reactor had a window of $2^1/_8$ by 11 inches and were of the cut-core type. These loops were separated and assembled with their d-c control coils in pairs around six of the twelve risers. The ends of the reactor risers were then formed over, collected into smaller bundles, and lugged for bolting to the input bus of their respective rectifier stacks. The transformer secondary leads thus became the 1-turn power winding of the reactors. The remaining six vertical risers formed the neutral leads of the 3-phase, half-wave connection. To keep the secondary inductance to a minimum, the neutral leads were brought up as close to their corresponding reactor lead as possible and then formed over to the neutral bus.

RECTIFIERS

Each assembly contained 12 selenium stacks, or two stacks per phase. Each stack contained 162 plates each $6^1/_4$ by $7^1/_4$ inches connected in parallel. These stacks were reported to be the largest of their type manufactured to date. The 12 stacks per unit were assembled into a common housing with six stacks (two per phase) on each side. Further to reduce the secondary inductance, the phase connections to six stacks were thoroughly transposed.

INTERPHASE

Electrically, the interphase reactor consisted of two turns, center-tapped. Mechanically, the interphase was used to provide the connection from the two neutral busses to the negative output terminal. Since each conductor of the reactor had to carry an average of 15,500 amp, it was necessary to utilize all of the $2^1/_8$- by 11-inch window of the reactor cores. The connection between the neutral bus and output terminal (which also became the windings of the interphase reactor) consisted of a pair of high-purity copper castings. These castings were provided with leaves on either end which simultaneously interleaved with the terminals of the neutral bus and the negative output terminal.

Because of the cost and bulk of the oil-to-air heat exchangers, the efficiency of the magnetic amplifier was of prime im-

Table II. Results of 1951 Magnetic-Amplifier Acceptance Tests

Type of Test	D-C Output	A-C Supply	Control	Response*	Efficiency
Minimum load test, two sections in parallel	2.0 volts, 28,700 amp	12.85 kv to 13.44 tap	−276 (a-t)		
Rated load test, three sections in parallel	13.9 volts, 93,360 amp, 1,340 kw	12.1 kv to 12.16 kv tap, 81.3% power factor, 1,620 kw	+720 a-t		82.8%
Nominal range test, four sections in parallel	3.7 volts, 47,200 amp	12.8 kv to 13.44 kv tap	− 76.8 a-t		
	6.75 volts, 78,600 amp	12.75 kv to 13.44 kv tap	0 a-t		
	11.5 volts, 105,800 amp	12.6 kv to 13.44 kv tap	− 60 a-t		
Double power response test, three sections in parallel	7.25 volts, 74,200 amp, 538 kw	12.1 kv to 13.44 kv tap	0 a-t	1.5 cycles	
	10.7 volts, 100,000 amp, 1,176 kw	12.1 kv to 13.44 kv tap	+137 a-t		
	7.25 volts, 74,200 amp, 538 kw	12.1 kv to 13.44 kv tap	0 a-t	1.5 cycles	

*See Fig. 7 for oscillographic traces of load voltage and current.

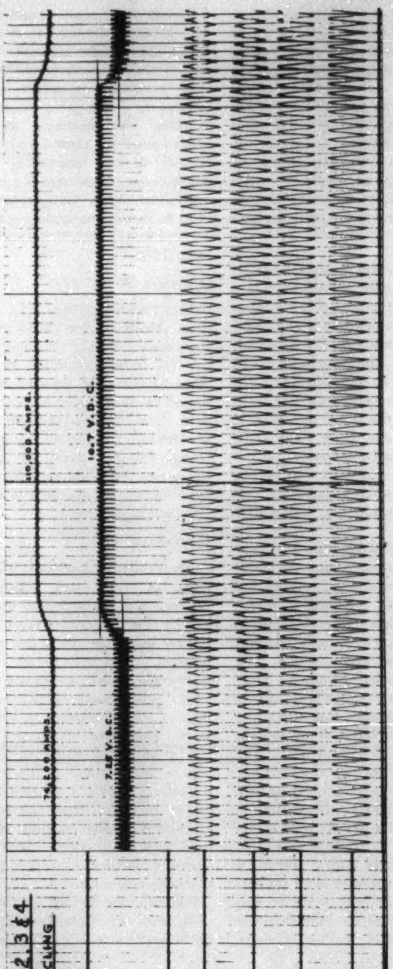


Fig. 7. Oscillographs of double power cycle

portance. Exclusive of auxiliary losses, the main losses in the amplifier consisted of selenium rectifier loss, transformer losses, and circuit *IR* losses. The stack drop was assumed to be the equivalent of 1.5 volts of applied alternating voltage. The total transformer losses (four units) at rated load and 14 volts direct current were computed to be 80 kw. The equivalent circuit d-c *IR* drop was computed at 0.2 volt.

$$\text{Computed magnetic amplifier efficiency} = \text{stack efficiency} \times \text{transformer efficiency} \times \text{d-c efficiency} = \frac{14.2\times0.855}{14.2\times0.855+1.5}\times\frac{1{,}950}{1{,}950+80}\times\frac{14}{14+0.2}=84.3\% \quad (13)$$

Test Data

Table II summarizes the pertinent results of the magnetic amplifier acceptance tests which were conducted late in 1951. The minimum load and rated load tests meet all specifications and, as was expected, the transformer primary taps were of assistance in meeting the extremes. The nominal range test was listed to show the output range readily available with 200 turns of bias winding and a single fixed transformer tap setting. In this case a change in control current from 348 ma minus to 300 ma positive swung the output voltage from 3.7 to 11.5 volts. For the double-power response test, the automatic cycling control switched the output voltage from 7.25 to 10.7 volts and the output kw from 538 to 1,176. The voltage and current oscillographs of the response test are shown in Fig. 7 where each vertical spacing represents 1 cycle. When the output voltage is switched from 7.25 toward 10.7 volts, the graph shows that the output voltage covers approximately 63% of the change in about 1.5 cycles. In switching from 10.7 to 7.25 volts, this is seen to be even less than 1.5 cycles. Although the oscillographs were taken for only three units, operation of the full magnetic amplifier would still give a response time of 2 cycles or less. Parallel operation of the units proved to be particularly successful. With loads of 120,000 amp, load division between units of ±1% was achieved in operation.

Although power gain was sacrificed for response in the double power test {the control supply voltage was 1,100 volts and the power gain $=[(1{,}176-538)1{,}000]/(1{,}100\times683)=1{,}023/1$}, the maximum possible power gain from the magnetic amplifier was substantial. For example, if the four 200-turn windings were connected

Vol. 7
Pt. 1
1955

in parallel, the total resistance would have been 15 ohms for three units. For the double-power test, the control power change would have been $0.685\times 15=7.03$ watts and the power gain would equal $638{,}000/7.03=90{,}500/1$.

Conclusion

The static magnetic amplifier described is, we believe, the largest of its type built to date. In this case, an application requiring the control of reasonably large blocks of power, which did not seem to lend itself to rotating magnetic amplifiers or other forms of power control, was satisfied with a static magnetic amplifier. The static magnetic amplifier met or exceeded all specifications. The speed of response of $1^{1}/_{2}$ to 2 cycles was well below that of comparable rotating amplifiers. The efficiency was high and the space requirements, in regard to the outdoor construction, were quite nominal. Maintenance and service have proved negligible.

The power application future of the static magnetic amplifier should be bright. Where dependability is at a premium, such as, e.g., in central generating stations, the static amplifier with its absence of moving and wearing parts should often replace the rotating amplifier. It is expected that with the demonstration that there are few size limitations upon static amplifiers, designers will find, that completely static excitation systems will further improve large alternator dependability.

The over 82% efficiency of the machine described is higher than can be achieved by other types of equipment at these low voltages. With selenium cells, static amplifier efficiencies of 80 to 85% can be expected with careful design at output potentials from 14 volts up to practically any desired voltage. At over 100 volts, the efficiency of the d-c static amplifier with selenium cells begins to be exceeded by other forms of equipment. However, this disadvantage of the static machine is only temporary. Silicon and germanium power rectifiers have now been developed to the point where engineers are designing large static magnetic amplifiers with efficiencies approaching 95%.

The advent of low-cost silicon and germanium power rectifier cells assures the extensive application of both d-c and a-c static magnetic power amplifiers. With the high gain of the self-saturating circuit, the amplifier reactor size becomes less than that of an autotransformer controlling comparable power. The new silicon or germanium power cells are considerably smaller than their controlling reactors. The static magnetic amplifier can therefore be manufactured with a smaller bulk and a cost that approaches that of a 2-winding transformer of the same rating. With high efficiency, extreme dependability, low weight, minimum bulk, and reasonable cost, the static magnetic amplifier will pre-empt much of the electrical field which are now served by other machines.

References

1. Electronics (book), Jacob Millman, Samuel Seely. McGraw-Hill Book Company, Inc., New York, N. Y., 1951.

2. Electronically Controlled Dry-Disk Rectifier, Allen Rosenstein, H. N. Barnett. *AIEE Transactions (Electrical Engineering)*, vol. 63, Jan. 1944, pp. 21–3; discussion p. 475.

3. Ground Power Supply For Modern Aircraft, T. B. Owen, J. A. Murray, Allen Rosenstein. *AIEE Transactions*, vol. 66, 1947, pp. 1297–1303.

4. Precision High-Current Computer Power Supplies, Allen B. Rosenstein. *AIEE Transactions*, vol. 73, pt. I, Sept. 1954, pp. 405–09.

5. On the Control of Magnetic Amplifiers, R. A. Ramey. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2124–28.

6. Applying Magnetic Amplifiers, W. J. Dornhoeffer, K. V. Krummenacher. *Electrical Manufacturing*, New York, N. Y., vol. 55, March–June 1951.

Analysis and Design of a Magnetic Frequency Multiplier

OTTO J. M. SMITH
MEMBER AIEE

JALAL T. SALIHI
STUDENT MEMBER AIEE

Synopsis: A magnetic frequency multiplier circuit operating from a 60-cycle-per-second (cps) 3-phase supply and delivering an approximately sinusoidal output at a frequency of 180 cps consisting of three identical nonlinear reactors of Deltamax cores inductively loaded is analyzed and a simple design procedure is outlined. The particular case of triple multiplication is chosen for convenience. The method of analysis may be extended to higher frequencies, either by using higher supply frequency or by using more phases. The power capability of the system depends upon the size of the cores, thus efficient multipliers for any desired power output may be designed.

IN EARLY days the need for higher frequencies for radio transmission aroused a considerable interest in frequency multiplication. The methods used were numerous. Broadly speaking, these may be divided into three classes, namely, machine, induction, and electronic apparatus. The machine type employ rotating parts while the induction type have the important advantage of being static where iron cores are electromagnetically excited in such a manner as to be periodically saturated, producing harmonics which are delivered to tuned loads. They are superior to electronic multipliers in that they use no tubes or rectifiers. Consequently, they are rugged, efficient, and can deliver large quantities of power.

During World War II, magnetic amplifiers became popular and extensive research on magnetic core materials was performed, yielding very high quality cores with almost rectangular hysteresis loop. The use of such cores as components for magnetic devices improves the performance and simplifies analysis and design. Thus using such cores, it is possible to build very efficient frequency multipliers having reliability and long life.

The circuit analyzed is similar to that of reference 1 where multipliers with tuned loads of power ratings ranging from 1 to 300 watts and with multiplication values ranging from 3 to 10 were mentioned. The analysis given in this reference went as far as deducing a relationship between applied voltage and firing angle. However the important fact that the relationship derived was based on the assumption of resistive loading was not pointed out. This relationship is gov-

---

Paper 55-69, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 11, 1954; made available for printing December 16, 1954.

Otto J. M. Smith and Jalal T. Salihi are with the University of California, Berkeley, Calif.

Portions of this paper have been abstracted from a thesis of Jalal T. Salihi. This thesis was submitted to the University of California in partial fulfillment of the requirement for the degree of Master of Science.

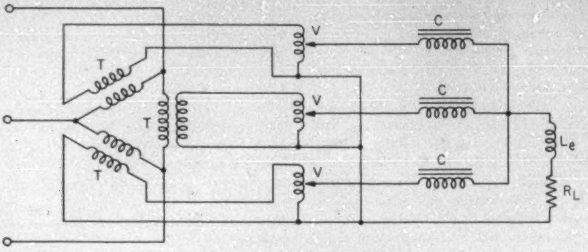


Fig. 1. Experimental multiplier circuit

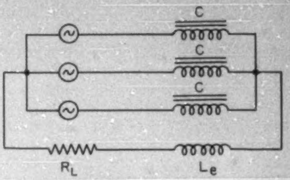


Fig. 3. Equivalent multiplier circuit

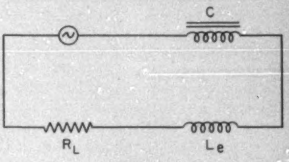
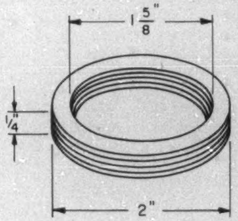


Fig. 4. Basic circuit for multiplier



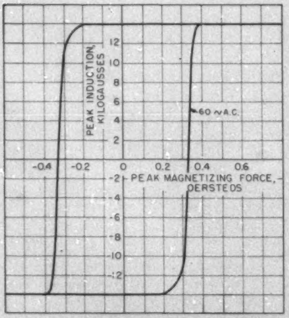
ASSUMED STACKING FACTOR FOR THE CORE = 0.79

Courtesy of Arnold Engineering Company

Fig. 5. Core of the experimental Deltamax reactor composed of stacked laminations of 0.001-inch thickness and wound with 5,000 turns of no. 33 wire gauge, having a resistance of 145 ohms

erned by the type of the load which must be specified as the first step towards analysis. In this paper the performance of a small unit for the cases of resistive and inductive loading is predicted mathematically and confirmed with measured results. The equations deduced are general and may be applied to multipliers of any size.

Equation 1 relates the conduction angle and applied voltage for the case of resistive loading and shows good correlation with practical measurement as seen in Fig. 8. However when this equation is converted to the appropriate form to be compared with the equation of reference 1 relating the firing angle and applied voltage the values do not check with the latter. The reason for this discrepancy is that the constant of integration corresponding to K_1, which represents the initial flux density at the instant of the start of the positive half-cycle of the supply voltage was taken to be zero; see Appendix I. This is not true since the reactor at this instant is still saturated in the negative direction caused by the preceding negative half-cycle of the supply voltage.



Courtesy of Arnold Engineering Company

Fig. 2. Magnetization curve of Deltamax

Circuit and Assumptions

The experimental circuit of Fig. 1 consists of three identical nonlinear reactors *C* with Deltamax cores of magnetization characteristics as shown in Fig. 2. Each reactor is fed in series with a common load consisting of the load resistance R_L and a smoothing linear reactor L_e from one of the phases of the 3-phase power supply through an isolation transformer *T* and a variac *V* providing the facility of changing the applied voltage. The equivalent circuit is shown in Fig. 3.

For the purpose of analysis the multiplier circuit is treated as three independent basic circuits, Fig. 4, each connected to the load separately and representing a branch of the multiplier circuit. The validity of this assumption follows from the fact that the nonlinear reactor of each basic circuit behaves in a manner similar in some respects to a gas tube. Thus, the current flows only for a portion of the cycle, the conduction angle depending upon the magnitude of the applied voltage. This assumption is not exactly true, however, for the experimental reactor of Fig. 5; the magnitude of the current which would flow outside the range of conduction is almost negligible, i.e., the impedance of the circuit to the flow of current is high compared with the conduction impedance. Consequently, since the applied voltages are displaced by 120 degrees, during the interval in which one of the basic circuits is conducting the remaining two are not, offering a very high impedance to the flow of current so that, when three of such circuits are connected to form a multiplier, the interaction between the branches is almost negligible and the multiplier output is simply equal to the sum of the outputs of the three circuits when each is connected to the load separately. This is true provided that the conduction angles for the basic circuits are equal to or less than 60 degrees of the supply voltage. The validity of this assumption may be seen from Fig. 6.

The assumptions made in analyzing the basic circuit of Fig. 4 are:

1. The winding resistance of the nonlinear reactor and the external linear reactor may be lumped with the load resistance.
2. The distributed and the winding capacitances of the reactors are negligible.
3. For high values of saturation current which apply to this case, the magnetization current, the eddy current, and the hysteresis losses are negligible and the magnetization characteristic may be represented, as shown in Fig. 7.

Nomenclature

A = effective core area, centimeters^2
B_c = critical value of flux density, gausses

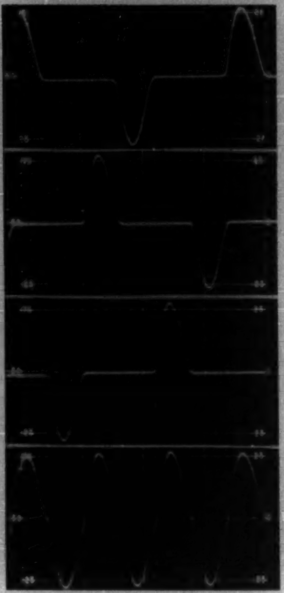


Fig. 6 (left). Oscillograms showing that the output of the multiplier is the sum of the outputs of three basic circuits each connected to the load separately. Multiplier composed of three experimental reactors of Fig. 5 inductively loaded; Q=0.5, L=0.34 henry

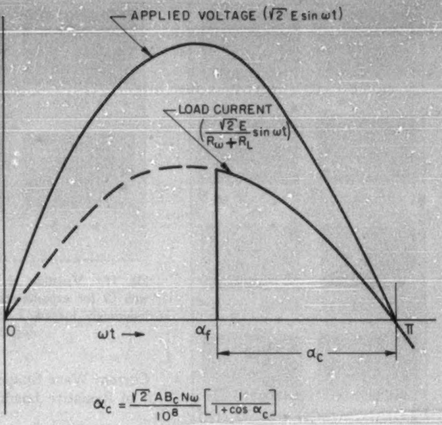

APPLIED VOLTAGE ($\sqrt{2}\,E\sin\omega t$)
LOAD CURRENT $\left(\frac{\sqrt{2}\,E}{R_\omega+R_L}\sin\omega t\right)$
0
ωt →
α_f
π
α_c
$\alpha_c=\frac{\sqrt{2}\,AB_cN\omega}{10^8}\left[\frac{1}{1+\cos\alpha_c}\right]$

Fig. 8 (right). Current for the case of resistive loading

l. 74
1
55

E=rms value of applied voltage, volts

E_{60}=rms value of the voltage applied to the basic circuit corresponding to a conduction angle of 60 degrees and a Q of 0.5, volts

$I_{max}=\frac{\sqrt{2}E}{(R^2+\omega^2L^2)^{1/2}}$

$I_{0\,max}$=maximum value of the output current for a conduction angle of 60 degrees and a Q of 0.5, amperes

I_0=rms value of the output current considered approximately sinusoidal for a conduction angle of 60 degrees and a Q of 0.5, amperes

i=instantaneous value of the current flowing in the basic circuit, amperes

L_s=saturation inductance of the nonlinear reactor winding, henrys

L_e=external or smoothing inductance, henrys

L=total inductance of the basic circuit at saturation=L_s+L_e, henrys

N=number of turns of the nonlinear reactor winding

P=power output of the multiplier corresponding to a conduction angle of 60 degrees and a Q of 0.5 for each of the basic circuits, watts

P_{max}=maximum power output of the multiplier corresponding to a conduction angle of 60 degrees and a Q of 0.5 for each of the basic circuits, watts

Q=effective Q of the basic circuit during saturation=$(\omega L)/R$

R_e=resistance of the external or smoothing inductance winding, ohms

R_L=load resistance, ohms

R_ω=resistance of nonlinear reactor winding, ohms

R=total resistance of the basic circuit=$R_\omega+R_L+R_e$, ohms

α_n=angle in the negative half-cycle of the a-c supply voltage wave through which the current caused by the positive half of the cycle continues to flow, degrees or radians

α_f=angle in a-c supply voltage wave at which firing occurs, degrees or radians

α_e=angle in a-c supply voltage wave at which extinction occurs, degrees or radians

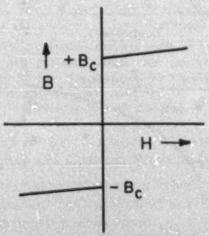


Fig. 7. Assumed magnetization characteristic; B_c=14,000 gausses

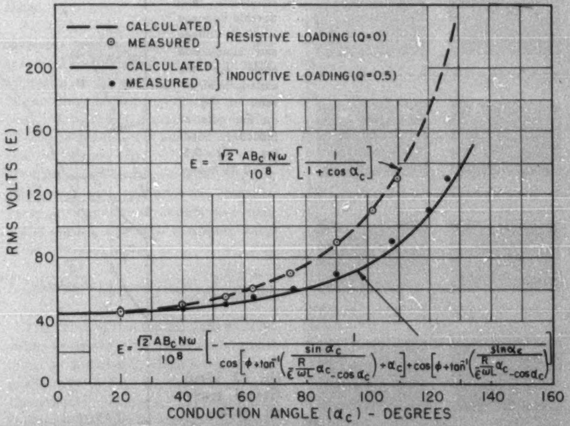


Fig. 9. Measured and calculated relationship between applied voltage and conduction angle for the experimental reactor of Fig. 5

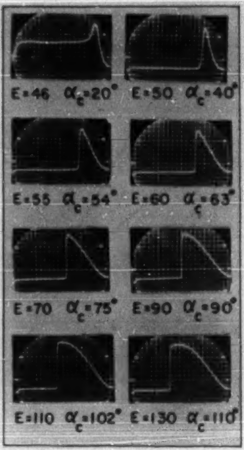


Fig. 10. Measured relationship between applied voltage and conduction angle for the experimental reactor of Fig. 5 for the case of resistive loading; R = 250 ohms

α_c = angle in a-c supply voltage wave through which conduction is maintained = $\alpha_e - \alpha_f$, degrees or radians
β = instantaneous value of the flux density, gausses
$\phi = \tan^{-1}(\omega L)/R$
Φ = total flux, lines
ω = angular velocity
η = efficiency

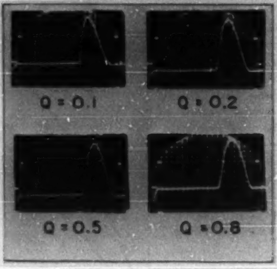


Fig. 11. Variation of current wave shape with Q for experimental reactor of Fig. 5 inductively loaded; L = 0.34 henry, α_c = 60 degrees

Current Wave Shape for the Case of Resistive Loading

The current flowing in the basic circuit of Fig. 4 for the case of resistive loading is a portion of a sinusoidal current which would flow when the saturable reactor is substituted by a resistance equal to its winding resistance. The current wave is shown in Fig. 8 and, to define it completely, the relationship between applied voltage and conduction angle must be known. This is derived in Appendix I and is

$$E = \frac{\sqrt{2} A B_c N \omega}{10^8} \left(\frac{1}{1 + \cos \alpha_c} \right) \tag{1}$$

To check the validity of this relationship, calculated values of E corresponding to different values of conduction angle are plotted in Fig. 9 together with measured results obtained from oscillograms of Fig. 10 which show good agreement.

The output of the tripler for the case of resistive loading contains only odd harmonics of three times the supply frequency. A Fourier analysis of the output is given in reference 2 where a general equation is derived by which the phase and amplitude of the harmonics corresponding to any value of conduction angle may be calculated.

Current Wave Shape for the Case of Inductive Loading

The output of the multiplier for the case of resistive loading is not sinusoidal and, as this requirement is essential for the output, several attempts have been made to attain sinusoidal approximation. Using a tuned output, the desired harmonic may be filtered; however, the addition of a smoothing inductance in series with the load is the simplest method to achieve this approximation. To estimate the value of the inductance necessary for an acceptable output wave shape,

Vol.
Pt.
195

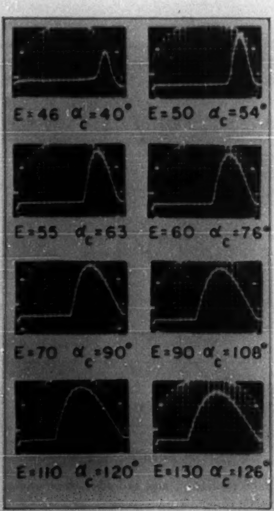
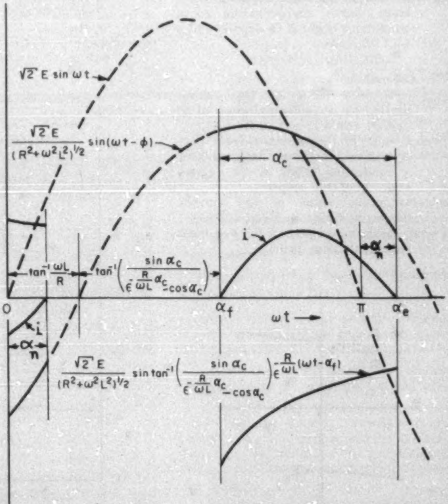


Fig. 12 (left). Measured relationship between applied voltage and conduction angle for the experimental reactor of Fig. 5 for the case of inductive loading, Q = 0.5, L = 0.34 henry


$\sqrt{2}\,E \sin \omega t$
$\frac{\sqrt{2}\,E}{(R^2+\omega^2 L^2)^{1/2}} \sin(\omega t - \phi)$
α_c
i
α_n
$\tan^{-1}\frac{\omega L}{R}$
$\tan^{-1}\left(\frac{\sin \alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c} - \cos \alpha_c}\right)$
0
α_f
ωt →
π
α_e
i
α_n
$\frac{\sqrt{2}\,E}{(R^2+\omega^2 L^2)^{1/2}} \sin \tan^{-1}\left(\frac{\sin \alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c} - \cos \alpha_c}\right)\epsilon^{-\frac{R}{\omega L}(\omega t - \alpha_f)}$

$$E = \frac{\sqrt{2} A B_c N \omega}{10^8}\left[-\frac{1}{\cos\left[\phi + \tan^{-1}\left(\frac{\sin \alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c} - \cos \alpha_c}\right) + \alpha_c\right] + \cos\left[\phi + \tan^{-1}\left(\frac{\sin \alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c} - \cos \alpha_c}\right)\right]}\right]$$

Fig. 13 (right). General graphical solution for current for the case of inductive loading

several oscillograms for the output current of the basic circuit were observed corresponding to different values of Q, while the conduction angle was fixed at 60 degrees by adjusting the applied voltage. These oscillograms are shown in Fig. 11, indicating that the output will approximate a sinusoid closely when Q is increased sufficiently. However, the addition of inductance will decrease the amplitude of the output. Thus, in choosing a suitable value for Q, a compromise should be made. In the authors' opinion, a convenient value to be chosen for Q is 0.5 which will be taken as the basis for the subsequent analysis.

The wave shape of the current corresponding to a given conduction angle is to be determined. This may be done in two steps: first, by deducing the relationship between applied voltage and conduction angle; and secondly, by deducing the relationship between this angle and the current.

RELATIONSHIP BETWEEN APPLIED VOLTAGE AND CONDUCTION ANGLE

This relationship is derived in Appendix I and is

$$E=\frac{\sqrt{2}AB_cN\omega}{10^8}\times$$

$$\left[-\frac{1}{\cos\left[\phi+\tan^{-1}\left(\dfrac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)+\alpha_c\right]+\cos\left[\phi+\tan^{-1}\left(\dfrac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)\right]}\right] \quad (2)$$

An examination of the relationship derived shows that for a fixed Q, changes in R and the corresponding changes in L do not affect the conduction angle.

To check the validity of equation 2, calculated values of E corresponding to different values of conduction angle are plotted in Fig. 9 together with measured results obtained from oscillograms of Fig. 12 which show good agreement. From the graphical relation of Fig. 9, the following conclusions may be made:

1. The critical voltage at which firing takes place is a characteristic of the reactor and is independent of load condition.
2. For a given applied voltage higher than the critical value, the presence of inductance in the load increases the conduction angle, thus making the reactor more sensitive to voltage changes.

RELATIONSHIP BETWEEN CONDUCTION ANGLE AND CURRENT WAVE SHAPE

The differential equation describing the operation of the basic circuit in the saturation region is of the form

$$L\frac{di}{dt}+R_i=\sqrt{2}E\sin\omega t$$

The solution of this differential equation is carried out in Appendix II. The expression derived for the current is

$$i=\frac{\sqrt{2}E}{(R^2+\omega^2L^2)^{1/2}}\times$$

$$\left[\sin(\omega t-\phi)-\sin\tan^{-1}\left(\frac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)e^{-\frac{R}{\omega L}\left[\omega t-\phi-\tan^{-1}\left(\frac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)\right]}\right] \quad (3)$$

Substituting for the value of E given by equation 2, a general expression for the current is

$$i=\frac{2AB_cN\omega}{10^8(R^2+\omega^2L^2)^{1/2}}\times$$

$$\left[-\frac{1}{\cos\left[\phi+\tan^{-1}\left(\dfrac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)+\alpha_c\right]+\cos\left[\phi+\tan^{-1}\left(\dfrac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)\right]}\right]\times$$

$$\left[\sin(\omega t-\phi)-\sin\tan^{-1}\left(\frac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)e^{-\frac{R}{\omega L}\left[\omega t-\phi-\tan^{-1}\left(\frac{\sin\alpha_c}{e^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)\right]}\right] \quad (4)$$

From this expression the wave shape of the current corresponding to any value of conduction angle may be determined showing that, for a fixed conduction angle corresponding to a fixed applied voltage, provided $Q=\omega L/R$ remains constant, the changes in L and the corresponding changes in R only affect the amplitude of the current but do not change the wave shape.

An analytical solution for the current involves tedious calculations. Using graphical methods and introducing some approximations, an expression for the current may be deduced readily. Equation 3 is represented graphically in Fig. 13 for the general case and in Fig. 14 for the particular case of interest where $Q=0.5$ and $\alpha_c=60$ degrees, indicating that the current wave shape for this case may be approximated fairly closely to a sinusoidal current having a maximum value of $I_{0\text{ max}}$ where

$$I_{0\text{ max}}=0.260\,I_{\text{max}},\ \text{and}\ I_{\text{max}}=\frac{\sqrt{2}E}{(R^2+\omega^2L^2)^{1/2}}$$

The value of the applied voltage E depends upon the conduction angle and is given by equation 2, which is represented graphically in Fig. 15, using $E/AN\omega$ as the ordinate. This graphical relationship is general and may be applied to any size of reactor with Deltamax cores, and the value of E corresponding to a conduction angle of 60 degrees may be found. Demoting this voltage by E_{60} and letting $\omega L=R/2$, the expression for the maximum value of the approximate sinusoidal current for any reactor reduces to

$$I_{0\text{ max}}=0.328E_{60}/R$$

and the rms value

$$I_0=0.232\,E_{60}/R \quad (5)$$

To study the effect of connecting three of the basic circuits to form a multiplier, oscillograms of the output wave shape for each circuit connected to the load separately were taken; in each case the conduction angle was adjusted to 60 degrees. Finally, the multiplier circuit was connected and a fourth oscillogram was taken. These oscillograms are shown in Fig. 6 which clearly indicates that the interaction between the branches is negligible; thus, the output of the multiplier is practically the sum of the outputs of the three circuits and is approximately sinusoidal.

Power and Efficiency

To fulfill an approximately sinusoidal output requirement the Q for each of the basic circuits and the conduction angle are fixed at 0.5 and 60 degrees respectively throughout; thus, the applied voltage is constant at E_{60}. For any changes in R, a corresponding change in the value of L is assumed to be made such that Q remains constant. Such changes will affect neither the firing angle nor the general wave shape of the output current; they merely alter the amplitude of the latter. Thus, the output current may be assumed approximately sinusoidal throughout these changes, the rms value of which is given by equation 5.

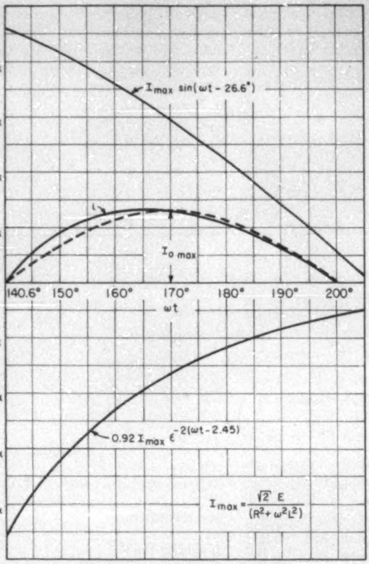


Fig. 14. Graphical determination of current i for the case of inductive loading, $Q=0.5, \alpha_0=60$ degrees

Assuming the resistance of the smoothing reactor to be a part of the load, the power output of the multiplier is

$$P=I_0^2R_L=0.0538\,\frac{E_{60}^2R_L}{(R_\omega+R_L)^2} \qquad (6)$$

From this equation it may be concluded that the maximum power $P_{\max}$ is delivered to a load having a resistance equal to the winding resistance of the reactor where

$$P_{\max}=0.0134\,\frac{E_{60}^2}{R_\omega} \qquad (7)$$

From equations 6 and 7

$$\frac{P}{P_{\max}}=\frac{4R_\omega R_L}{(R_\omega+R_L)^2}$$

For high values of load current the presaturation current is negligible and the expression for efficiency may be written as

$$\eta=\frac{R_L}{R_\omega+R_L}\times 100 \text{ per cent}$$

Values of $P/P_{\max}$ and η corresponding to different values of R_L are plotted in Fig. 16. It must be pointed out that these relations are only true for the limited range of operation where the presaturation current is negligible compared with the saturation current.

Design and Calculations

Given three identical Deltamax cores, the windings may be designed as follows:

1. From Fig. 15, the value of $E/AN\omega$ corresponding to 60 degrees may be found.
2. Deciding on a suitable value for E_{60} and knowing ω and A, the number of turns N may be calculated.
3. From tables a suitable wire size may be chosen to fill the available window area, and the resistance of the windings may be found.
4. From efficiency and power curves of Fig. 16, a suitable operating point may be chosen so that the value of the load resistance may be calculated.
5. The inductance of the winding at saturation is usually less than the value required for a $Q=0.5$, thus sufficient external inductance may be added to fulfill this requirement.
6. Making use of equation 7 and Figs. 15 and 16, the output power of the multiplier may be calculated.

The process of design may be reversed

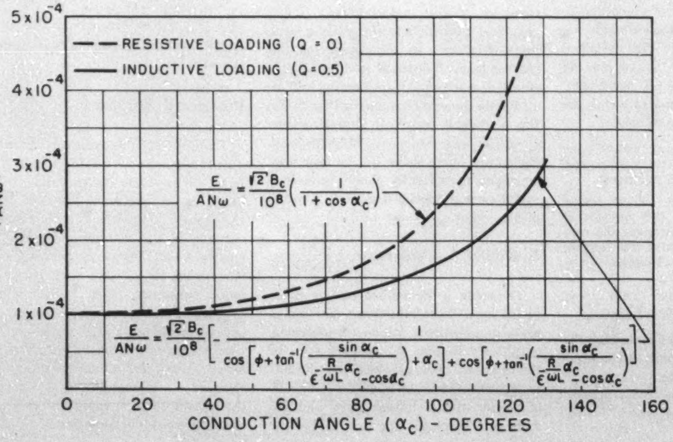


Fig. 15. Relation between E/ANω and conduction angle for reactors with Deltamax cores, $B_c=14{,}000$ gausses

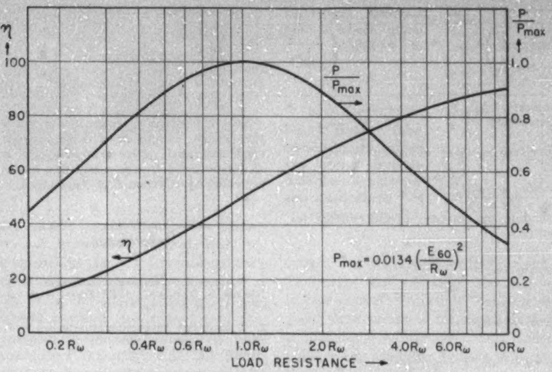


Fig. 16. Variation of power and efficiency with load

and multipliers for any given power output may be designed by the choice of an appropriate size of cores.

For a multiplier using three of the experimental reactors, the maximum power is $0.0134\times(52)^2/145=0.25$ watt compared with a measured power of 0.22 watt.

The experimental reactors used were already wound for a different purpose where the winding was occupying less than one-third of the available window area; consequently, the 0.25 watt does not represent the full power capability for the size of the core. An investigation of the power output must be based on a winding making full use of the window area of the experimental core which may be designed by following the procedure suggested. The maximum power output for such a winding is estimated as 0.7 watt.

For more power, more weight is required. The relationship between weight and power output may be deduced by applying equation 7 to reactors of different sizes. This was not investigated; however, it was observed that if n identical reactors of negligible winding reactances are connected in parallel and a voltage is applied, the conduction angle is the same as that for a single reactor having the same applied voltage, and the combination is equivalent to a single reactor having a winding resistance equal to R_w/n. Knowing this, and making use of equation 7, it may be concluded that the maximum power output for a multiplier having n identical reactors per branch is n times greater than that of a multiplier using a single reactor per branch, thus the power may be increased linearly as the weight.

Conclusions

The output of the multiplier may be considered as the sum of the outputs of three independent basic circuits. For any of the basic circuits, the critical voltage at which firing takes place is a characteristic of the reactor unaffected by the load conditions while for voltages higher than this value, the presence of the inductance in the load increases the conduction angle, thus making the reactor more sensitive to changes in voltage. The addition of inductance also improves the output wave shape which may be made approximately sinusoidal if enough inductance is added. However, the addition of inductance in series with the load will decrease the amplitude of the output, thus a compromise between power output and wave shape must be made. Following the design procedure outlined, frequency multipliers for any desired power ratings and efficiency may be built.

Appendix I. Relationship Between Applied Voltage and Conduction Angle

The differential equation describing the operation of the basic circuit of Fig. 4 in the presaturation region is

$$N\frac{d\Phi}{dt}\times10^{-8}+L_e\frac{di}{dt}+Ri=\sqrt{2}E\sin\omega t$$

Neglecting the $Ri+L_e di/dt$ drop, the differential equation reduces to

$$N\frac{d\Phi}{dt}\times10^{-8}=\sqrt{2}E\sin\omega t$$

Since $d\Phi=Ad\beta$

$$d\beta=\frac{\sqrt{2}E\times10^8}{AN}\sin\omega t$$

Integrating

$$\beta=K_t+\frac{\sqrt{2}E\times10^8}{AN\omega}\int_{t_1}^{t}\sin\omega t d\omega t$$

The foregoing equation simply says that the instantaneous value of the flux density is equal to the initial flux density plus a varying flux density proportional to the volt-time area of the applied voltage.

Case of Resistive Loading

For this case, the extinction is at an angle of π radians, as shown in Fig. 8, thus the angle at which the voltage cycle beginning at $\omega t=0$ starts to contribute the volt-time area necessary to raise the flux is at $\omega t=0$.

Letting $K_t=-B_c$, then

$$\beta=-B_c+\frac{\sqrt{2}E\times10^8}{AN\omega}\int_0^{\omega t}\sin\omega t d\omega t$$

At the instant of firing $\beta=+B_c$ and $\omega t=\alpha_f$.

$$B_c=-B_c+\frac{\sqrt{2}E\times10^8}{AN\omega}\int_0^{\alpha_f}\sin\omega t d\omega t$$

$$=-B_c+\frac{\sqrt{2}E\times10^8}{AN\omega}(1-\cos\alpha_f)$$

Solving for E yields

$$E=\frac{\sqrt{2}AB_cN\omega}{10^8}\left(\frac{1}{1-\cos\alpha_f}\right)$$

Since $\alpha_f=\pi-\alpha_c$

$$E=\frac{\sqrt{2}AB_cN\omega}{10^8}\left(\frac{1}{1+\cos\alpha_c}\right) \qquad (8)$$

Case of Inductive Loading

Because of the lagging effect introduced by the inductance, the current continues to flow beyond the point where the supply voltage drops to zero at the end of the positive half-cycle. Thus, the angle at which the voltage cycle beginning at $\omega t=0$ starts to contribute the volt-time area necessary to raise the flux is at the angle of $\alpha_n=\alpha_f+\alpha_c-\pi$, as shown in Fig. 13.

Letting $K_t=-B_c$, then

$$\beta=-B_c+\frac{\sqrt{2}E\times10^8}{AN\omega}\int_{\alpha_f+\alpha_c-\pi}^{\alpha_f}\sin\omega t d\omega t$$

At the instant of firing $\beta=+B_c$ and $\omega t=\alpha_f$. Therefore

$$B_c=-B_c+\frac{\sqrt{2}E\times10^8}{AN\omega}\int_{\alpha_f+\alpha_c-\pi}^{\alpha_f}\sin\omega t d\omega t$$

$$=-B_c-\frac{\sqrt{2}E\times10^8}{AN\omega}\times[\cos(\alpha_f+\alpha_c)+\cos\alpha_f]$$

Solving for E yields

$$E=\frac{\sqrt{2}AB_cN\omega}{10^8}\left(-\frac{1}{\cos(\alpha_f+\alpha_c)+\cos\alpha_f}\right)$$

Substituting for the value of α_f in terms of α_c deduced in Appendix II gives

$$E=\frac{\sqrt{2}AB_cN\omega}{10^8}\left[-\frac{1}{\cos\left[\phi+\tan^{-1}\left(\dfrac{\sin\alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)+\alpha_c\right]+\cos\left[\phi+\tan^{-1}\left(\dfrac{\sin\alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)\right]}\right] \quad (9)$$

For any fixed value of $\omega L/R$ the foregoing equation relates the rms applied volts to the conduction angle.

Appendix II. Relationship Between Conduction Angle and Current for the Case of Inductive Loading

The differential equation describing the operation of the basic circuit of Fig. 4 at saturation is

$$L\frac{di}{dt}+Ri=\sqrt{2}E\sin\omega t$$

Solving this equation and substituting the initial condition $i=0$ at $\omega t=\alpha_f$, the expression deduced for the current is

$$i=\frac{\sqrt{2}E}{(R^2+\omega^2L^2)^{1/2}}\times\left(\sin(\omega t-\phi)-\sin(\alpha_f-\phi)\epsilon^{-\frac{R}{\omega L}\alpha_c(\omega t-\alpha_f)}\right)$$

Substituting the second boundary condition $i=0$ at $\omega t=\alpha_f+\alpha_c$, shown in Fig. 13, in the foregoing equation gives

$$\sin[\alpha_c+(\alpha_f-\phi)]=\sin(\alpha_f-\phi)\epsilon^{-\frac{R}{\omega L}\alpha_c}$$

Expanding

$$\sin\alpha_c\cos(\alpha_f-\phi)+\cos\alpha_c\sin(\alpha_f-\phi)=\sin(\alpha_f-\phi)\epsilon^{-\frac{R}{\omega L}\alpha_c}$$

Dividing by $\cos(\alpha_f-\phi)$ and solving for $\tan(\alpha_f-\phi)$

$$\tan(\alpha_f-\phi)=\frac{\sin\alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}$$

From which,

$$\alpha_f=\phi+\tan^{-1}\left(\frac{\sin\alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)$$

This equation relates the firing angle to the conduction angle for any value of $Q=\omega L/R$.

Substituting α_f in the equation for current gives

$$i=\frac{\sqrt{2}E}{(R^2+\omega^2L^2)^{\frac{1}{2}}}\times\left[\sin(\omega t-\phi)-\sin\tan^{-1}\left(\frac{\sin\alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)\epsilon^{-\frac{R}{\omega L}\left[\omega t-\phi-\tan^{-1}\left(\frac{\sin\alpha_c}{\epsilon^{-\frac{R}{\omega L}\alpha_c}-\cos\alpha_c}\right)\right]}\right] \quad (10)$$

This equation is a general solution for the current corresponding to any given conduction angle where the value of E is also determined by the conduction angles expressed by equation 2 of Appendix I.

References

1. MAGNETIC FREQUENCY MULTIPLIER, L. J. Johnson, S. E. Rauch. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 448–52.

2. A MAGNETIC FREQUENCY MULTIPLIER, J. T. Salihi. *M.S. Thesis*, University of California, Berkeley, Calif., July 1954.

3. RADIO FREQUENCY CHANGERS, A. N. Goldsmith. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 3, Mar. 1915, pp. 55–79.

4. FREQUENCY MULTIPLICATION, PRINCIPLE AND PRACTICAL APPLICATION OF FERRO-MAGNETIC METHODS, N. Lindenblad, W. W. Brown. *AIEE Transactions*, vol. 44, Apr. 1925, pp. 491–96.

5. SELF-SATURATION IN MAGNETIC AMPLIFIERS, W. J. Dornhoefer. *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 835–46.

Properties of Silicon Power Rectifiers

E. F. LOSCO
NONMEMBER AIEE

THE STATIC electrical properties of small-area silicon *P-N* junction rectifiers prepared by the alloy fusion process have been described by Pearson and Sawyer[1] in 1952. Since that time rectifiers of the same type and having about the same *P-N* junction area (in the order of 0.0005 to 0.001 square centimeter (cm)) have become commercially available. The development of these rectifiers confirms theoretical expectations as to the superior electrical properties of silicon over germanium for semiconductor devices. These superior properties, which include extremely low reverse currents, high rectification ratios, and the ability to operate successfully at high ambient temperatures, are attributed to the wider energy gap in silicon (1.15 electron volts) as compared to germanium (0.75 electron volt).

Because the silicon rectifiers now available commercially are quite small in *P-N* junction area, they are necessarily limited to applications involving relatively low-level currents of 30 to 50 milliamperes (ma). The rms power outputs of these units, when used as single-phase, half-wave rectifiers with resistive load, are confined to 1 watt or less. It is well recognized therefore that the development of silicon rectifiers capable of handling several watts, or even several hundred watts of rectified power, would represent an important forward advance and open up many new power applications.

This paper is concerned with a description of some experimental silicon rectifiers capable of handling power outputs in the order of 500 to 1,200 watts. Both static and dynamic electrical properties to be described indicate that it is entirely feasible to realize such high-power handling ability for relatively large area (0.05 square cm) *P-N* junction silicon rectifiers prepared by the alloy fusion process.

Preparation of Silicon Power Rectifiers

The silicon power rectifiers have a junction area of approximately 0.05 square cm, or about 50 times the junction area of familiar commercially available units. The junctions are prepared by the alloy fusion process employing *N*-

---

Paper 55-152, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 25, 1954; made available for printing November 29, 1954.

E. F. Losco is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

The author wishes to acknowledge the considerable technical assistance of Dr. S. J. Angello in the solution of difficult basic problems. Acknowledgment is also made to C. S. Duncan for supplying single crystals of silicon, to P. Demmler for preparing the P-N junctions, to T. E. Ebert for obtaining data on d-c characteristics, and to H. R. Camp for carrying out the a-c power tests.

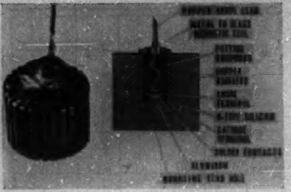


Fig. 1 (left). Construction of silicon power rectifier

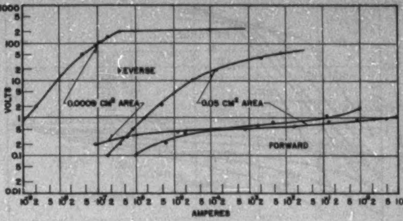


Fig. 2 (right). Comparison of d-c reverse and forward characteristics of small- and large-area fused silicon P-N junction rectifiers

type single crystal silicon. The silicon crystals, grown from the melt with the long axis in the [111] crystallographic direction, have electrical resistivity in the order of 1 to 5 ohm-cm. While several elements in Group III of the Periodic System have been employed for fusing the *P-N* junction, it has been found that aluminum is the most practical *P*-type impurity element to use. The melting point of aluminum, 660 degrees centigrade (C), is very convenient for alloy fusion. In addition, the low vapor pressure of aluminum at temperatures considerably above its melting point permits the fusion operation to be carried out without difficulty either in vacuum or inert gas atmospheres.

As illustrated schematically in Fig. 1, the rectifiers are assembled with suitable anode and cathode terminals which protect the brittle silicon wafer from damage during handling and permit easy assembly of electrical connections. Also illustrated in Fig. 1 is the method of encapsulation in a cylindrical copper heat exchanger having vertical fins arranged symmetrically around its outside periphery. The rectifier is mounted at the base of a blind cylindrical hole within the copper radiator, which serves also as one electrode of the rectifying cell. The other electrode is a wire lead emerging from the top of the radiator through a metal-to-glass bushing, which hermetically seals the rectifier. A suitable potting agent fills the cavity between rectifier and top of the heat exchanger, and electrically insulates the rectifier terminals. The completely assembled unit, shown in Fig. 1, is 1 inch in outside diameter and about 1 inch high.

D-C Characteristics at Room Temperature

The d-c reverse and forward characteristics of a 0.05 cm² fused *P-N* junction silicon rectifier is compared in Fig. 2 with those of a similar but very small area, 0.0009-cm² fused *P-N* junction unit. It will be noticed that the small area unit has very much lower leakage current. At 1 volt, for example, the leakage current is only slightly over 10^{-9} ampere. The larger area power rectifier has a leakage current of about 10^{-6} ampere or about three orders of magnitude higher at the same voltage level. At higher voltages this difference persists with not much change in the slopes of the two curves until reverse voltage breakdown is approached.

The reason for the higher leakage currents observed in the large-area rectifiers is not entirely clear. It is suspected, however, that the presence of shunts around the junction periphery may be responsible. The longer periphery and the greater chance of incomplete removal of shunts by chemical etching following the fusion operation suggests this possibility. Recent unpublished data on the effectiveness of various etchants on reducing the surface recombination velocity on germanium rectifiers indicates that surface preparation of the single crystal wafer is of extreme importance. This, no doubt, is also true in the case of silicon units and it is expected that future developments in this field will materially reduce leakage currents.

Neither the small- nor the large-area silicon rectifiers in Fig. 2 display a true reverse saturation current which would be expected from the theoretical current voltage relationship given by

$$I = I_0[\exp(qV/kT) - 1]$$

where

I = current, amperes per cm²
I_0 = a constant determined by silicon properties (such as lifetime of minority carriers, mobility, and resistivity)
q = electronic charge
V = voltage across the junction (plus for forward and minus for reverse currents)
k = Boltzmann's constant
T = absolute temperature

The reason for lack of current saturation in the reverse direction has been attributed to the low lifetime of minority current carriers in silicon. According to McAfee and Pearson,[2] working on junctions grown from the melt, the saturation current due to the minority carriers' diffusing into the space charge region is less than the reverse current due to carriers generated by traps within the space charge region, the trap reverse current being proportional to the width of the space charge region. Pearson and Sawyer[1] later extended this work and indicated lack of saturation also in small-area fused junctions.

Since the large-area fused junctions display similar reverse characteristics, it appears that improvement in this direction must come from the preparation of more nearly perfect single-crystal silicon having higher lifetimes, in the order of 100 to 300 microseconds instead of the at present available 10 to 50 microseconds. Occasionally, fused junction silicon rectifiers having reasonably good reverse saturation currents are prepared. This is more common in small-area rather than in large-area junctions. In such cases it is felt that a portion of the silicon ingot from which the crystal was prepared was exceptionally free of lattice imperfections.

Another interesting difference between the reverse d-c characteristics of small- and large-area fused junction silicon

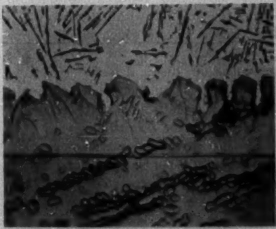


Fig. 3. Fused P-N junction in N-type silicon at 500 times magnification. Horizontal line is P-N junction separating original N-type silicon (bottom) from recrystallized P-type silicon. Aluminum-silicon eutectic is above recrystallized silicon. Lineage defects are evident within the silicon

L. 74
1
55

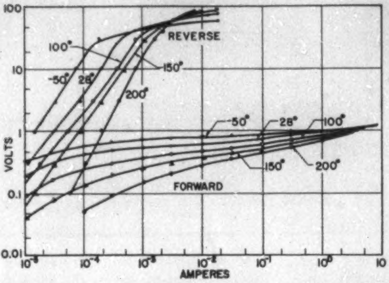
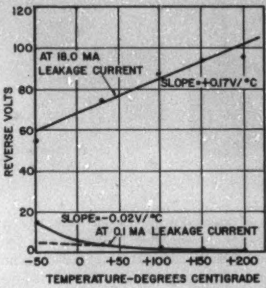


Fig. 4 (left). Effect of temperature on d-c reverse and forward characteristics of silicon power rectifier

Fig. 5 (right). Effect of temperature on d-c reverse voltage characteristics at two levels of leakage currents



rectifiers is the sharpness of the breakdown voltage. The small area junctions have a sharp knee beyond which the voltage remains essentially constant over several orders of magnitude of reverse current. The larger area junctions, on the other hand, do not show a sharp breakdown but display a more gradual bend in the curve which is often referred to as a "soft knee." The softness varies somewhat from unit to unit and is often improved by chemical etching of the junction's periphery. Etching, however, does not completely overcome softness.

One possible reason for the inability to realize sharp breakdown voltages in large-area fused junctions may be associated with presence of imperfections in the silicon single crystal. A photomicrograph of a cross-sectioned silicon rectifier, 0.05 cm² in junction area, is illustrated in Fig. 3 at 500 times magnification. The original *N*-type silicon crystal, at the bottom of the photomicrograph, is seen to have numerous lattice defects believed to be lineage (exaggerated in size as a result of chemical etching). These defects pass through the *P-N* junction (which appears as a horizontal line) and continue to develop above the junction within the recrystallized P-type silicon, which has the same crystal orientation as the *N*-type. The presence of these defects, especially along the *P-N* junction plane no doubt prevent full realization of theoretically expected junction properties.

The d-c forward characteristics of the small-area rectifier indicates a forward current of 0.1 ampere at 1 volt. This corresponds to a current density of 110 amperes per cm². The rectification ratio at 1 volt is about 100,000,000, reflecting the very low leakage currents obtainable in small *P-N* junctions. The slope of the forward characteristic curve, however, increases relatively fast above about 1 volt. This appears to be characteristic of small-area silicon rectifiers.

The larger area fused junction silicon rectifiers have better forward characteristics than the smaller area units. At 1 volt, for example, the forward current is 10 amperes, which corresponds to a current density of 200 amperes per cm² for the 0.05-cm² area. In addition, the slope of the forward characteristic curve does not increase as rapidly over 1 volt as does that for the small-area junction. This is very desirable since the bulk of the heat developed in the junction results from the forward rather than the reverse characteristics. For power applications in which the dissipation of heat becomes a serious problem, the ability to realize low forward voltage drops is quite important. The large-area units also display a fairly high rectification ratio. Fig. 2 indicates a rectification ratio of about 5,900,000 at 1 volt. Values of this order can be realized in the 0.05-cm² *P-N* junctions, despite the relatively high reverse leakage, because the forward current density at 1 volt is improved considerably.

D-C Characteristics as Affected by Temperature

Temperature affects the reverse leakage of the 0.05-cm² silicon rectifiers very appreciably. This is shown graphically in Fig. 4 in which the d-c characteristics of a representative rectifier at −50, 28, 100, 150, and 200 C are plotted with the use of log-log scale. At low values of voltage, it is seen that leakage current increases as test temperature increases. It is also evident that the slopes of the reverse curves do not change markedly as test temperature is changed. When the applied voltage is sufficiently high, so that breakdown conditions are approached, the curves cross and leakage values become lower, the higher the test temperature. This effect has been previously described for small-area silicon junctions.[1]

Fig. 5 illustrates graphically the magnitude of reverse voltage change as affected by temperature at relatively high and low values of leakage currents. At an arbitrary leakage of 18 ma (beyond the knee of the current versus voltage curves in Fig. 4) the reverse voltage increases about 0.17 volts per C rise in temperature. At 0.1 ma leakage current, the voltage versus temperature curve has a negative slope and the voltage decreases at the rate of 0.02 volts per C over the straight-line portion of the curve. The magnitude of these changes may vary from unit to unit; the values given are therefore indicative of the trend for the 0.05-cm² area junctions described.

Fig. 4 also illustrates the effect of temperature on forward d-c characteristics. In this case, the effect is straightforward in that the forward voltage drop is lower, the higher the test temperature. At low values of forward current, temperature changes result in larger differences in forward drop then at high values of forward current. For a forward drop

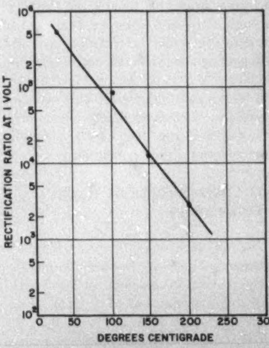
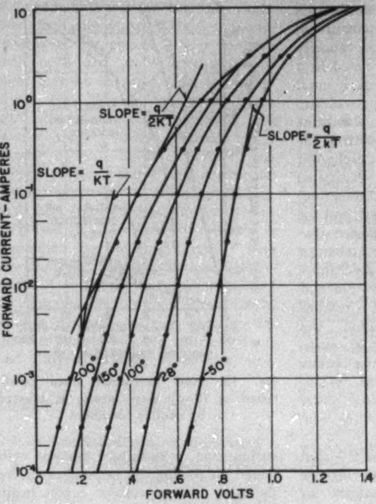


Fig. 6. Rectification ratio at 1 volt as affected by temperature



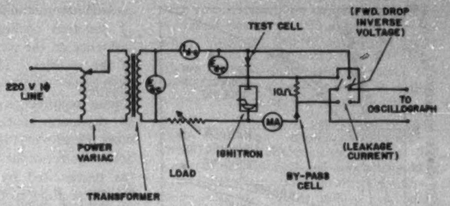


Fig. 7 (left). Approach of forward characteristics of power rectifiers to the current-voltage expression, $I = I_0$ [exp $(qV/kT)-1$], at various temperatures

Fig. 8 (above). Single-phase half-wave circuit diagram for power tests

of 0.5 volt for example, the forward current changes from 4.5×10^{-5} amperes at −50 C to 2×10^{-1} amperes at 200 C, a change of almost four orders of magnitude. At 1 volt, the corresponding change is from 1.5 to 3 amperes, or less than one order of magnitude.

The effect of temperature on rectification ratio for the application of 1 volt in both the forward and reverse directions is indicated in Fig. 6 for a typical 0.05-cm² area silicon rectifier. At room temperature, the forward to reverse current ratio is in the order of 500,000. This value decreases to 13,000 at 150 C and to 3,000 at 200 C. The increase in leakage current is primarily responsible for the drop in rectification ratio. Rectification ratios much higher than these, in the order of 6,000,000 at room temperature have been obtained as was previously shown in Fig. 2. Variations in the slope of the curve for rectification ratio as temperature is increased have also been noted. Uniformity of properties, in this respect, leave something to be desired. It is felt, however, that spread of results for the most part is dependent on material uniformity and fusion techniques. Improvements in the latter are expected materially to raise average rectification ratios.

It has been found that the current-voltage expression previously indicated is correct for forward currents of small area silicon *P-N* junctions. Fig. 7 illustrates the closeness of approach to this expression for the 0.05-cm² junctions at various temperatures. If close agreement is expected, then a plot of the forward current on a logarithmic scale versus forward voltage on an arithmetic scale should yield curves the straight-line portions of which have slopes corresponding to q/kT. It is apparent from Fig. 7 that at high temperatures, 200 C, there are two straight-line portions to the curve. One of these, which occurs at relatively low voltages (up to about 0.15 volt), has a slope which corresponds to q/kT as predicted. The other, occurring at higher voltages (0.2 to 0.5 volt), has a slope which corresponds to $q/2kT$.

At very low temperatures, −50 C, the straight-line portion of the curve falls in the range of 0.7 to 0.9 volt and its slope corresponds to $q/2kT$. It is possible that extension of this curve to lower values of voltages would also show a straight-line portion with slope q/kT. For intermediate temperatures, the straight-line portion of the curves has slopes which lie between q/kT and $q/2kT$ for the voltage range investigated. Here again it might be expected that the forward characteristics would follow the current-voltage expression at lower values of voltage.

A-C Characteristics

The a-c power tests carried out on the 0.05-cm² area silicon rectifiers involved 60-cycle single-phase half-wave circuits with resistive loads. The test circuit is illustrated in Fig. 8. The current-carrying ability of the encapsulated rectifiers was determined by noting how the peak values of forward voltage drop and leakage current varied as the load current was increased. This was done both for the case of natural convection air cooling and for the case of forced convection of air over the external fins of the copper heat exchanger. As the tests progressed, measurements were taken of the temperature of the heat exchanger about $^1/_{16}$ inch below the base terminal of the rectifier.

Results obtained for a rectifier operated at a peak inverse voltage of 100 volts are shown in Fig. 9. For natural air convection cooling, the forward drop and leakage curves vary almost linearly with load current until the latter exceeds a 5- to 6-ampere average. At 6 amperes, the leakage current has reached a value of 10 ma and is rising sharply, more so than the forward drop. It appears, in fact, that continued increase in the load current much beyond 7 amperes, at which point the leakage current is about 20 ma, would result in a runaway condition that would lead to failure of the rectifier. Therefore, for these particular conditions the silicon power rectifier can safely handle a load current of 6.5 amperes average at a 100-volt peak inverse. This corresponds to 10.2 amperes rms which, for the 0.05-cm² area junction, indicates a current density of 200 amperes rms per cm². The rms power output for natural air cooling,

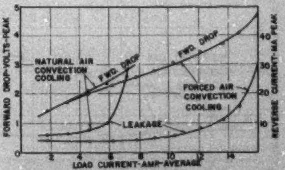


Fig. 9. A-c forward and reverse characteristics when power tested at 100-volt peak inverse for two methods of cooling

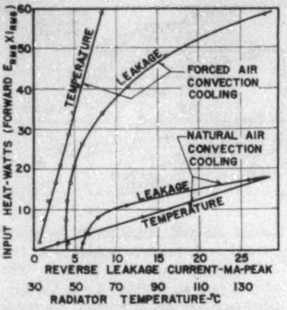


Fig. 10. Heat dissipation when power tested at 100-volt peak inverse as affected by cooling conditions

then, is $I_{rms} \times E_{rms} = 10.2 \times 50$, or 510 watts. The d-c power output is 6.5×32, or 207 watts.

For forced air convection cooling, the slope of the forward drop curve is somewhat lower than before and remains nearly linear until a load current in the neighborhood of a 15-ampere average is reached. The peak leakage current is reduced considerably, as would be expected, by air cooling. Its initial value is appreciably reduced, the curve having essentially a zero slope until quite high values of load current are reached. To reach a 20-ma peak leakage current, the load current must be more than doubled (from 7 amperes for natural air cooling to 15.25 amperes for forced air cooling). If a load current of 15 amperes average is taken as being a safe upper limit for forced air convection cooling, we can say that the encapsulated rectifier under these conditions can safely handle 24 amperes rms or about 480 amperes per cm². The rms power output of the rectifier then corresponds to 24×50 or 1,200 watts. The d-c power output in this case is 15×32 or 480 watts.

As illustrated in Fig. 9, the effectiveness of junction cooling has a pronounced influence on the power-handling capacity of the rectifier. The copper-finned heat exchanger used for the experimental silicon rectifiers is limited in its ability to dissipate heat. This is illustrated more clearly in Fig. 10 by plotting the actual power input to the heat exchanger (as represented by the heat dissipated by the rectifier in the forward direction) against the peak leakage current and the exchanger temperature. It is seen that the upper limit of the heat exchanger is approached when the curve for leakage current approaches zero slope. For natural air cooling, this condition is reached at about 20 watts' dissipation. For forced air cooling, it is over 60 watts. Peak leakage currents at these values are in the order of 25 to 30 ma. Thus, the previously mentioned output power ratings of the rectifier are on the conservative side.

It is also evident from Fig. 10 that heat exchanger temperatures as high as 140 C were reached for natural air cooling. At these temperatures, there was no indication of instability of the rectifier despite the relatively high leakage current. When forced air cooling is employed, the maximum exchanger temperature was only about 65 C. With more efficiently designed heat exchangers, and more effective cooling, more heat can be dissipated and advantage can be taken of the inherently high power-handling capacity of the rectifier.

Figs. 11 and 12 are the same types of plots shown in the previous two figures. The curves, however, are for a different encapsulated rectifier tested at 70-volt instead of 100-volt peak inverse. At the lower voltage higher forward currents can be handled since leakage currents are reduced. Thus, Fig. 11 shows that an 8.5-ampere average with natural air cooling and a 28-ampere average with forced air cooling can be applied before the peak leakage current curves begin to rise rapidly above 10 ma. These values correspond to current densities of 280 and 880 amperes rms per cm². The output power is 470 and 1,540 watts rms, or 190 and 625 watts d-c respectively. The heat dissipation, shown in Fig. 12, is in the order of 12 and 68 watts for the two cooling conditions. This corresponds to rectifier efficiencies of about 98 and 96 per cent respectively. It is also evident from Fig. 12 that the maximum temperature reached by the heat exchanger when natural air cooling was employed was 230 C. The temperature of the *P-N* rectifying junction itself is slightly higher than this, indicating the ability of silicon rectifiers to operate at theoretically predicted high ambient temperatures.

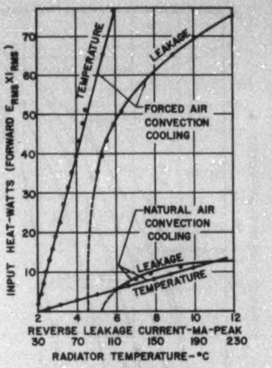


Fig. 12. Heat dissipation when power tested at 70-volt peak inverse as affected by cooling conditions

Numerous experimental silicon recti-

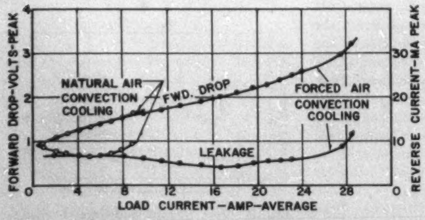


Fig. 11. A-c forward and reverse characteristics when power tested at 70-volt peak inverse for two methods of cooling

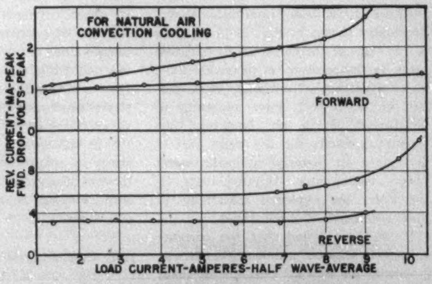


Fig. 13. Range of a-c forward and reverse characteristics for several rectifiers power tested at 50-volt peak inverse

fiers with 0.05-cm^2 junction area have been made and tested. The a-c characteristics so far presented were for two such units and are representative of what may be expected from individual units. Fig. 13, on the other hand, illustrates the range for forward and reverse characteristics obtained for a number of units tested at 50-volt peak inverse using natural air convection cooling only. Curves of this type along with those shown in Figs. 10 and 12 have been found to be extremely valuable in evaluating dynamic properties. Safe power ratings can thus be easily determined by noting the load current required to cause forward drop or, more likely, leakage current to rise sharply for different values of peak inverse voltage and different methods of junction cooling.

Summary and Conclusions

Experimental silicon rectifiers with a *P-N* junction area of approximately 0.05 cm^2 have been prepared by the alloy fusion process using *N*-type silicon and aluminum. These units are hermetically sealed in copper heat exchangers equipped with external vertical fins for either natural or forced air convection cooling. The units were subjected to both d-c and a-c tests over a series of temperatures varying from −50 to +230 C. From these tests the following conclusions are drawn.

D-C PROPERTIES

1. The rectifiers have quite low forward voltage drops, about 0.8 volt at 1 ampere, and can handle currents as high as 200 amperes d-c per cm^2 of junction area. In this respect, they compare favorably with very small area (0.001-cm^2) silicon-fused junction diodes.

2. They do not display a low saturation current, or a sharp breakdown voltage, as do very-small-area junctions. For many power applications, however, the leakage currents are low enough to be entirely adequate.

3. At room temperature, rectification ratios as high as 6,000,000 have been obtained at 1 volt. The ratio decreases as temperature increases, but usable values are obtained up to at least 200 C. A rise in leakage current is primarily responsible for the drop in rectification ratio with increasing temperature.

4. D-c breakdown voltages increase with increasing temperature. The forward d-c characteristics appear to follow the theoretical current-voltage expression, $I=I_0[\exp(qV/kT)-1]$ for low values of forward voltage which are obtained at high temperatures. For high voltages and low temperatures, the junctions follow more closely the expression $I=I_0[\exp(qV/2kT)-1]$.

A-C PROPERTIES

1. With only natural air convection cooling, the encapsulated rectifiers, arranged in a single-phase half-wave rectifier circuit with resistive load, are capable of handling a current density of 200 amperes rms per cm^2. They can deliver an output power in the order of 500 watts rms (207 watts d-c). For 100 volts peak inverse, this corresponds to a load current of a 6$^1/_2$-ampere average or 10.2 amperes rms.

2. With forced air convection cooling, the rectifier's power-handling ability is more than doubled. Current densities of 480 amperes rms per cm^2 and output power of over 1,200 watts rms (480 watts d-c) have been achieved. These values correspond with load currents of a 15-ampere average or 24 amperes rms at 100-volt peak inverse.

3. The silicon rectifiers themselves are capable of handling higher powers than those indicated. Power is limited now by the relatively low efficiency of the heat exchanger being used. With natural air convection cooling the upper limit of heat dissipation of the heat exchanger is in the order of 20 watts. For forced air convection cooling, heat dissipation is increased to the order of 70 watts.

4. Silicon rectifiers have been successfully operated with the use of natural air convection cooling with heat exchanger temperatures up to 230 C and with *P-N* junction temperatures somewhat higher than 230 C.

References

1. SILICON P-N JUNCTION ALLOY DIODES, G. L. Pearson, B. Sawyer. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 40, no. 11, Nov. 1952, pp. 1348–51.

2. THE ELECTRICAL PROPERTIES OF SILICON P-N JUNCTIONS GROWN FROM THE MELT, K. B. McAfee, G. L. Pearson. *Physical Review*, New York, N. Y., vol. 87, July 1952, p. 190.

Junction Transistors Used as Switches

R. L. BRIGHT
ASSOCIATE MEMBER AIEE

MUCH literature has been devoted to the use of junction transistors as linear amplifiers but there has been very little information presented on the signal properties of these devices. As large signal devices, transistors have unique characteristics which enable them to perform many functions which were impossible with previously known devices. It is the purpose of this paper to point out some of these characteristics and to give a few illustrative applications.

The characteristics of a NPN fused junction transistor will be introduced by considering its behavior in the circuit shown in Fig. 1(A). The d-c volt-ampere characteristics for the input and output terminals are shown in Figs. 1(B) and (C). The curve in (B) shows that the input circuit looks much the same as a simple diode; the base current is very small if the base voltage is negative and becomes relatively large as the base voltage is made positive. The curve of (C) shows that, if the input is reversed biased 1/10th volt or more so that a small negative base current flows, then the load current in (C) will be very small and the transistor is said to be cutoff *A*. If the input is open-circuited, the base current will be zero and the load current will increase to the value at *B*. As the base voltage is made more positive, the operating point will move through the so-called linear region *C*. If the base current is raised to the value of 30 milliamperes (ma), the operating point will be at *D* and a further increase in the base-driving current will have practically no effect on the load current. This condition of operation is defined as saturation and is characterized by a very small collector-to-emitter voltage drop.

If the driving voltage V_b is abruptly switched from a negative value to a positive value, high enough to ensure saturation, the output circuit will be switched from *A* where the load current is very small to *D* where the load current is large and the voltage drop V_{ce} is very small. Under this condition, the transistor behaves similarly to the polarized relay shown in Fig. 2.

High-Level Operation

One advantage of operating the transistor in this on-off manner rather than as a linear amplifier is that the transistor

---

Paper 55-156, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing December 16, 1954.

R. L. BRIGHT is with the Westinghouse Electric Corporation, Pittsburgh, Pa.

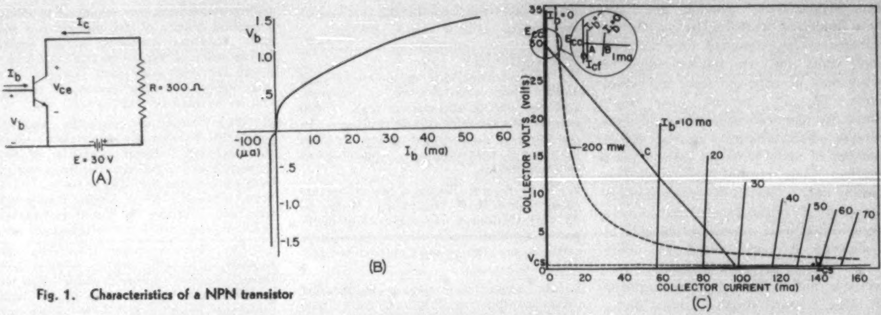


Fig. 1. Characteristics of a NPN transistor

power dissipation for such switching operation is only a small fraction of the peak load power that can be controlled whereas the dissipation is equal to or greater than the load power when linear operation is employed. For example, in the circuit of Fig. 1 the transistor dissipation at cutoff is equal to the base voltage times the base current plus the product of the collector current and voltage equals 1 volt × 2 microamperes + 30 volts × 30 microamperes = 0.9 megawatt (mw). At the saturation the transistor dissipation is $1.1\times30+0.25\times100=59$ mw and the load power is 3 watts. Thus, the load power is many times the dissipation. This fact means that the junction transistor is a very efficient device for controlling power. At present, it is limited to circuits having voltages on the order of 50 to 100 volts but, when used as a switch in such circuits, its efficiency is far higher than either gaseous or high vacuum tubes.

Note that the dissipation corresponding to C in Fig. 1 is 15 volts × 50 ma = 750 mw and is thus equal to about 1/4 of the peak load power. Since in practice the operating point must move continuously from A to D, the instantaneous power dissipation during this transition will be quite large and, therefore, the duration of the transition period must be small enough so that the total energy dissipated does not overheat the unit. This requirement is ordinarily easy to satisfy at frequencies up to 1,000 cycles per second; but at the high audio-frequencies the transition time often becomes an appreciable part of the cycle and this restriction may become serious. If the load current of Fig. 1 is assumed to have the wave shape shown in Fig. 3, the transition dissipation averaged over the entire cycle is P_L peak $(fT/3)$ where P_L = peak load power $=E^2/R$.

A typical value for T is 5 microseconds and thus, if the frequency f is 400 cycles per second, the average transition power dissipation is only $400.5\times10^{-6}/3=0.06$ per cent of the peak load power. However, at a frequency of 40 kc, it represents 6 per cent of the load power and may be the dominant term contributing to the total dissipation.

Referring again to Fig. 1(C), it may be seen that the amount of base current required to saturate the transistor depends upon the maximum load current. If the value of E/R is 100 ma, then a base current of 30 ma will be sufficient to ensure saturation. If E/R is 150 ma, a base current of at least 70 ma is necessary. At low-current densities, the relation between the minimum driving current required for saturation and the load current is nearly linear but, at the high-current densities often employed in power control circuits, this relationship becomes very complicated and must generally be determined experimentally. For high-gain units at low currents, the required base current may be as low as 1/100th of the load current while at high-current densities, this factor may be as high as 1/2. Even though the current gain at these high currents is only 1/2, the power gain (defined as the load power divided by the base input power) may be quite large. For example, if $E=30$ volts and $I_b=70$ ma, the curves of Fig. 1 show that $I_c=150$ ma, $V_b=1.7$, thus, $P_b=120$ mw. $P_L=4.5$ watts, so that the power gain equals 4,500/120 equals 37.5. The minimum required base drive P_b is completely determined if the maximum value I_c is specified and is independent of E if E/R is kept constant; therefore, the power gain at any particular load current is directly proportional to the source voltage E.

Fig. 4 shows a set of experimental curves giving the following quantities as a function of the load current I_L for the common emitter connection of Fig. 1. The curve of Fig. 4(A) shows the minimum base current required for saturation; (B) shows the base input power $P_b=V_bI_b$, the collector-emitter power $P_c=V_{ce}I_c$, and the total sum or total power dissipation P_c; and (C) shows the power gain per volt of the supply voltage $=I_L/P_b$. The per-cent power loss per volt of the supply voltage $=I_L/P_c$. (This is a measure of the over-all efficiency.)

Such curves furnish a convenient basis for a comparison between the various methods of connecting the transistor in the circuit. Curves of the total power dissipation P_c versus I_c are shown in Fig. 5 for the three basic connections: common emitter, common base, and common collector. These curves show that the common collector connection has the highest efficiency, but it has a disadvantage in that the base must be

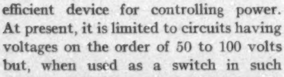


Fig. 2 (left). Equivalent polarized relay

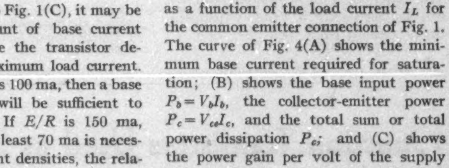
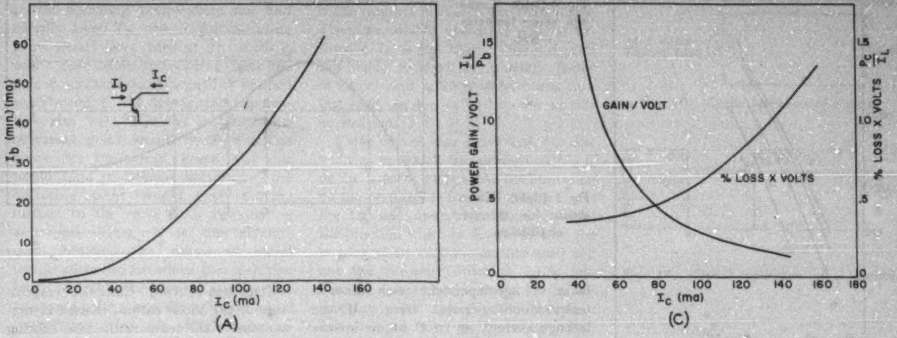


Fig. 3 (right). Wave shape showing transition interval



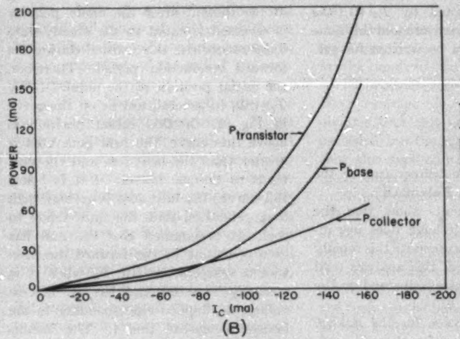


Fig. 4. High-level performance curves

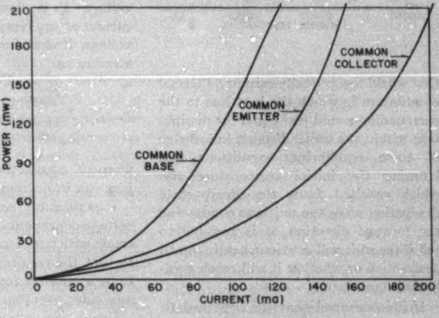


Fig. 5. Comparison of basic connections

negative with respect to the emitter at cutoff and, therefore, the input voltage at cutoff must be greater than the supply voltage E while during conduction it is only 1 to 2 volts. This is often awkward circuitwise and, therefore, the common emitter connection seems preferable for most high-power applications. For the benefit of those familiar with the ordinary transistor rating sheets, it should also be mentioned that in switching operation the breakdown potentials for the common emitter and common base connections are the same, it is only for linear operation that the voltage rating of the common emitter is less than that for common base.

Maximum Ratings

The power-handling capability of a given transistor can be determined if the cooling coefficient, the room temperature leakage, the ambient temperature, the duty cycle, and the amplitude and wave forms of the voltage that the unit is required to block are all known. The maximum load-handling capability may be limited by any one of the three following factors: the melting point of the terminals; the breakdown voltage of the junctions; or the phenomenon of thermal runaway. Thermal runaway is a consequence of the fact that the leakage current increases so rapidly with temperature that a condition may exist where a small rise in temperature results in an increase in leakage and, hence, an increase in power dissipation which is greater than the additional heat transfer to the surroundings resulting from this temperature rise. When this situation exists, the unit becomes hotter and hotter until it is destroyed. The melting point of the terminal is usually the limiting factor when the inverse voltages do not exceed about 10 volts. The junction breakdown voltage is a direct limitation only at very low ambient temperatures. For most applications, thermal runaway is the determining factor.

This thermal runaway criterion will be analyzed in detail for: a junction diode that has been conducting long enough to reach steady state whereupon the applied voltage is suddenly reversed; a junction diode in cyclic operation where the period is small compared to the thermal time constants; and a transistor in either cyclic or single on-off service.

A more complete statement of the first situation is: Assume that a diode is operating in a prescribed ambient temperature and has been conducting a specified constant current in the forward direction for a period long enough for the unit to reach thermal steady state. The polarity is then suddenly reversed and the unit is required to block a constant voltage. How does the maximum inverse voltage that can be applied without runaway depend upon the ambient temperature and the value of the dissipation that existed prior to the reversal?

The polarity of thermal runaway can be analyzed by determining the conditions

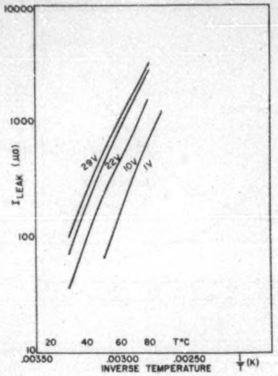


Fig. 6 (left). Leakage versus temperature

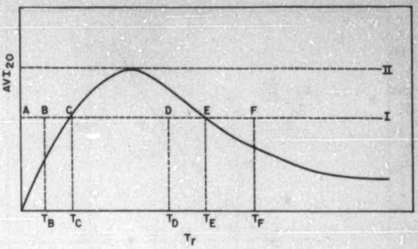


Fig. 7 (right). Condition for thermal equilibrium

that would theoretically exist at a thermal equilibrium in which the heat loss to the surrounding would just equal the dissipation within the unit. From a knowledge of these equilibrium conditions and knowing the initial temperature rise which resulted from the steady-state dissipation while the unit was conducting the forward direction, it is possible to tell if the unit will continue heating until destruction or whether it will reach equilibrium condition.

If it is assumed that the heat loss to the surroundings is directly proportional to the temperature rise above ambient, then a necessary condition for thermal equilibrium is that the temperature rise T_r, in degrees centigrade (C), is equal to the cooling coefficient A, in C per mw, times the dissipation within the unit. If the diode is blocking, this dissipation is equal to the inverse voltage times the leakage current and, since this leakage is a function of the temperature rise, its functional form must be investigated before the equilibrium can be computed.

Theoretically, the leakage current of a junction diode is an exponential function of the temperature of the form exp $(-8{,}350/T)$ for germanium and exp $(-12{,}900/T)$ for silicon and is independent of the applied inverse voltage above about 1/4 volt. In practice, this temperature dependence is followed quite closely at a low inverse voltage but the current is experimentally found to vary considerably with the applied voltage. In the remainder of this analysis, it will be assumed that this voltage-dependent component of the leakage current varies with temperature in the same manner as the theoretical constant component so that the leakage current can be written as a voltage-dependent term times a temperature-dependent term. If the leakage current at 20 C at an inverse voltage V is denoted by $I_{20}(V)$, the current at any temperature and this same voltage V can then be written for germanium as

$$I = I_{20}(V)\frac{\epsilon^{\frac{8{,}350}{-T(°K)}}}{\epsilon^{\frac{8{,}350}{-293}}}$$

where T equals the temperature of the junction in degrees Kelvin (K).

Experimental results show that the voltage-dependent leakage does not increase with temperature quite this rapidly so that the results of this analysis tend to be somewhat conservative and on the safe side; see Fig. 6.

Using this expression for the leakage current, the equation for thermal equilibrium can be written as

$$T_r = A\,VI = A\,VI_{20}(V)\frac{\epsilon^{\frac{8{,}350}{-T}}}{\epsilon^{\frac{8{,}350}{-293}}}$$

Letting

$$T = T_a + T_r + 273$$

where T_a = ambient temperature in C and rearranging

$$A\,VI_{20}(V) = \frac{T_r\,\epsilon^{\frac{8{,}350}{-293}}}{\epsilon^{\frac{-8{,}350}{273+T_a+T_r}}}$$

Thus, if the applied voltage and the ambient temperature are known, $I_{20}(V)$ and A can be measured and the equilibrium temperature rise T_r can be evaluated. A graph of $A\,VI_{20}(V)$ versus T_r for an ambient of 40 C is illustrated in Fig. 7.

This curve represents a locus of operating points at which the steady-state internally generated power is equal to the heat loss and, from it, the steady-state temperature rise that a diode would assume for any given value of inverse voltage can be read. Since the temperature of the diode cannot change instantaneously, the temperature rise existing at the instant after the diode polarity is reversed is equal to the steady-state temperature rise that existed during the forward conduction period. Therefore, the initial position of the point $(A\,VI_{20}, T_r)$ will, in general, not lie on the curve in Fig. 7. If this initial position is above this curve, the heat generated is greater than the heat loss and the unit tends to become hotter; if it is below this curve, the heat loss is greater than that generated and the unit tends to cool. It is assumed that the unit has been operating in the forward direction with a temperature rise T_B when it is suddenly reversed and has an inverse voltage V_1 applied corresponding to the horizontal dashed line I. The instantaneous operating point will then be at B which is above the steady-state curve and the unit will heat up until it reaches thermal equilibrium at a temperature T_C corresponding to point C.

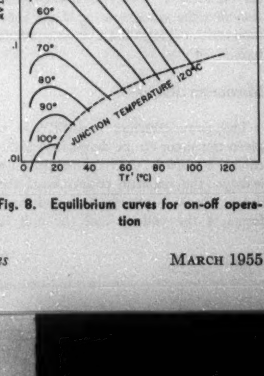


Fig. 8. Equilibrium curves for on-off operation

Vol.
Pt. 1
1955

If the operating temperature rise had initially been T_D, the initial operating point would have been at D, which is below the steady-state curve, and the unit would tend to cool until it reached equilibrium C. If the initial temperature rise were T_F, the instantaneous operating point would be F which is above the curve and, hence, the unit would tend to become hotter and the operating point would keep moving further to the right until the unit is destroyed. Thus, E is the dividing point between the values of initial temperature rise for which the operation tends toward the finite steady-state operating point C and the values of initial temperature rise for which runaway occurs. Note also that runaway will occur for all values of voltage which are higher than the value corresponding to line II regardless of the initial temperature rise.

Fig. 8 is a family of such curves for various ambient temperatures. If the inverse voltage V is specified and A and $I_{20}(V)$ are measured, it is possible to use this set of curves to determine the maximum forward dissipation that is permissible at any given ambient. Conversely, a test limit may be imposed in $I_{20}(V)$ that will insure that the unit will not run away with prescribed voltage, dissipation, and ambient temperature. For example, assume that it is desired to operate a diode at an ambient temperature of 60 C with a forward dissipation corresponding to a temperature rise of 40 C and an inverse voltage of 50 volts. The curves indicate that runaway will be avoided when the polarity reverses, if the product AVI_{20} is less than 0.088. Therefore, stable operation will be assured if I_{20} is less than $0.088/50A$, and this value reduced by a safety factor of 20 per cent or so may be taken as a test limit for diodes to be used in this application.

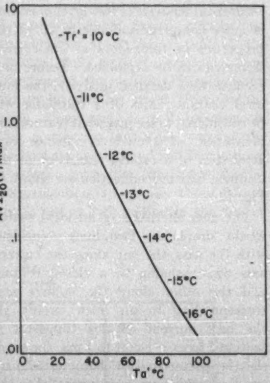


Fig. 9 Equilibrium curves for cyclic operation

These curves also show that, for low voltages at which the product AVI_{20} is on the order of 0.02 or less, the limitation on the permissible forward dissipation is the top-safe operating temperature of the junction which is determined by the melting point of the elements used and not the runaway criterion. For germanium devices having indium terminals (melting point 145 C), the maximum safe temperature is probably about 120 C.

The limitation imposed by the junction breakdown voltage is indirectly included in the analysis in that the value of $I_{20}(V)$ becomes very large as the breakdown voltage is approached. It is obvious that this voltage can never be exceeded, but the analysis has also shown that the maximum permissible voltage under high ambient conditions may be much less than the junction breakdown value.

The more common case in which the diode carries forward current for a fraction of each cycle and then sustains a reverse voltage for the remainder of the cycle can be handled by an extension of this analysis. Let

P_f = average power dissipation in the conducting or forward direction
P_i = average power dissipation in the inverse direction
A_f = cooling coefficient for the conduction period
A_i = cooling coefficient for the inverse period

The averaging is carried out over the entire cycle, e.g., the average forward power P_f is equal to the energy dissipated in the forward direction per cycle divided by the entire period. It is assumed that the frequency is high compared to the reciprocal of the thermal time constant of the diode.

The temperature rise is then given by

$$T_r = A_iP_i + A_fP_f$$

where

$$A_iP_i = A_i(I_iV_i)_{av} = A_iVI_{20}(V)_{av}\frac{\epsilon^{\frac{-8,350}{T_a+T_r+273}}}{\epsilon^{\frac{8,350}{-298}}}$$

$$AVI_{20}(V)_{av} = (T_r - A_fP_f)\frac{\epsilon^{\frac{8,350}{-298}}}{\epsilon^{\frac{-8,350}{T_a+T_r+273}}}$$

Substituting:

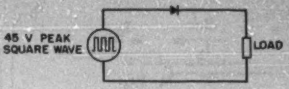


Fig. 10. Cyclic operation of a diode

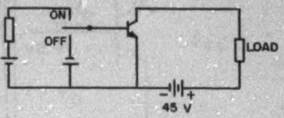


Fig. 11. On-off operation of a transistor

$$T_a' = T_a + A_fP_f$$
$$T_r' = T_r - A_fP_f$$

we obtain

$$A_iVI_{20}(V)_{av} = T_r'\frac{\epsilon^{\frac{8,350}{-298}}}{\epsilon^{\frac{-8,350}{T_a'+T_r'+273°}}}$$

This expression has the identical form as the corresponding expression for the steady-state solution of a diode with a constant inverse V, as given previously, where $VI_{20}(V)_{av}$ replaces $VI_{20}(V)$ and T_a' and T_r' replace T_a and T_r.

In other words, for cyclic operation, the temperature rise caused by the forward dissipation effects the reverse characteristic in exactly the same manner as an equivalent increase in the ambient. The general curves of Fig. 8 are thus applicable to this case also where the effective ambient temperature is the true ambient plus the temperature rise caused by the forward dissipation and the abscissa is the temperature rise above this effective ambient. Actually, for cyclic operation, the only point of interest on these curves is the value of the maximum voltage that can be sustained for any given effective ambient and this information can be replotted as the single curve of Fig. 9. As an example, consider the circuit shown in Fig. 10.

If the constants of the diode are $A_i = A_f = 0.15$ C per mw and $I_{20}(45) = 0.020$ ma, what is the maximum forward dissipation permissible at an ambient of 40 C? Since this is a 50-per-cent duty cycle

$$AVI_{20}(V)_{av} = 0.15\frac{(45)(0.020)}{2} = 0.0675$$

From Fig. 9, the highest effective ambient at which this value of $AVI_{20}(V)_{av}$ is permissible is 74 C. Hence, the maximum temperature rise caused by the forward dissipation must not exceed

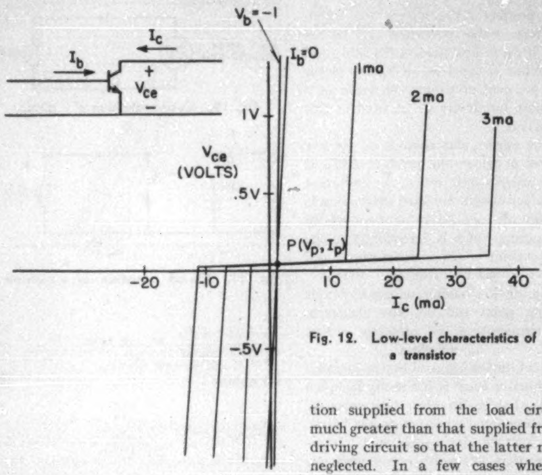


Fig. 12. Low-level characteristics of a transistor

74 to 40=34 degrees which corresponds to about 440 mw peak (again 50 per cent duty cycle). If the source voltage were sinusoidal, the form of solution would be the same except that value of $VI_{20}(V)_{av}$ would have to be found graphically from an oscilloscope.

In on-off duty, as in the first example, this diode would have a value of $AVI_{20}(V)=0.135$ at 45 volts and, from Fig. 8, could tolerate a forward temperature rise of 60 C at a 40-degree ambient, which corresponds to a 400-mw forward dissipation. Thus, the peak forward power in cyclic operation can safely exceed the maximum allowable in on-off operation so long as the period of oscillation is small compared to the thermal time constant. In general, a single on-off operation is the most severe for both diodes and transistors.

In extending this analysis to transistors used as switches, the only change required is that the power dissipation supplied by both the driving and load circuit must be taken into account. Thus, the temperature rise during conduction is A times the sum of these powers. Usually, the driving power is of the same order as the dissipation supplied from the load circuit and, hence, both must be computed.

In the cutoff conduction, the term $AVI_{20}(V)$ must be generalized to include a term $AV_dI_{20}(V_d)$ for the drive circuit plus a term $AV_1I_{20}(V_1)$ for the load circuit. In most cases, the load supply voltage is much higher than the driving cutoff voltage and the leakage dissipation supplied from the load circuit is much greater than that supplied from the driving circuit so that the latter may be neglected. In a few cases where the transistor is being used in the inverted connection and the saturation current of the emitter junction is very flat, this assumption may not be valid and both powers may have to be included.

As an example, for the circuit of Fig. 11, determine the maximum total forward dissipation permissible for a single on-off operation at an ambient of 40 C with a load supply voltage of 45 volts and a cutoff leakage current in the load circuit of 20 microamperes and a cutoff leakage current of 2 microamperes at 1.5 volts in the basic circuit at 20 C.

The total leakage power at 20 C is

$$1.5(0.002)+45(0.02)=0.9 \text{ mw}$$

Thus

$$AI_{20}(V)V=0.15+0.9=0.135$$

From Fig. 8, the maximum permissible forward power dissipation is

$$\frac{0.60 \text{ C}}{0.15 \text{ C per mw}}=400 \text{ mw}$$

Note the highest ambient temperature at which this transistor could sustain 45 volts can be found from either Fig. 8 or 9 to be 63 degrees and the maximum forward dissipation at this temperature is about

$$\frac{13.5 \text{ C}}{0.15}=90 \text{ mw}$$

Establishing ratings for transistors in class-A and B operation is more difficult. The same thermal equilibrium condition $T_r=AP$ still holds, but the temperature dependence of P can be very involved. The collector current, e.g., is I_e+I_{c0} where I_{c0} has the same type of exponential dependence previously given for the leakage current. Most amplifier circuits incorporate stabilizing resistors in series with the emitter so that a change in the quiescent current feeds a voltage change back to the base so as to change the emitter current; thus, the emitter current changes with temperature in a manner dependent upon the circuit used. Similarly, the collector voltage depends upon the total current and, hence, also varies in a complicated fashion with temperature. Therefore, the expression of P as a function of temperature depends upon the circuitry and no general runaway criterion can be given.

The ratings established here only guarantee that the transistor will not burn out at the upper temperature limits; they do not guarantee that circuits incorporating these transistors will operate satisfactorily. At high temperatures, the leakage currents may be an order of magnitude or more greater than at room temperature and the circuit must be designed to perform correctly even with this high leakage. For both transistors and diodes, it was the conducting dissipation and not the conducting current that was specified. The relation between dissipation and current may be obtained from curves of the type shown in Fig. 5.

Low-Level Operation

In addition to having much higher efficiencies than tubes, transistors also have the advantage that they can operate at much lower voltages: in many cases the transistor behaves as a nearly ideal switch down to potentials on the order of 100 microvolts. The theoretical analysis of low-level operation is quite complete and nearly all of the properties of transistors at low-current densities can be explained. Before proceeding to a detailed analysis, the low-level characteristics of a transistor will be considered to see just what features are of interest. The low-level collector characteristics of a typical transistor in the common emitter connection are shown in Fig. 12.

The characteristics of an ideal switch would simply be two lines coincident with the axis, the one along the current axis corresponding to a closed contact and the other along the voltage axis corresponding to an open circuit. If the base current of this transistor is switched from −1 volt to 1 ma, the small signal collector-emitter characteristics are two lines which are nearly parallel to the axis but which do not intersect at the

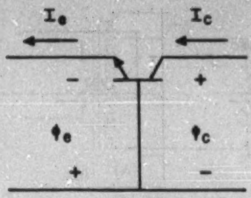


Fig. 13. General transistor equations

$$I_e = I_{e0}\frac{\left(\epsilon^{\frac{q\phi_e}{kT}}-1\right)-\alpha_i I_{c0}\left(\epsilon^{\frac{q\phi_c}{kT}}-1\right)}{1-\alpha_i\alpha_n}$$

$$I_c = \frac{+\alpha_n I_{e0}\left(\epsilon^{\frac{q\phi_e}{kT}}-1\right)-I_{c0}\left(\epsilon^{\frac{q\phi_c}{kT}}-1\right)}{1-\alpha_i\alpha_n}$$

$$\alpha_i I_{c0} = \alpha_n I_{e0}$$

where

α_n = current gain in the normal connection and is ordinarily between 0.90 and 0.999
α_i = current gain in the inverted connection and is ordinarily between 0.3 and 0.6, although it may be about 0.98 in special symmetrical units
I_{c0} = saturation current of the collector junction
I_{e0} = saturation current of the emitter junction

origin but rather at a point $P(I_p, V_p)$ in the first quadrant. In other words, the transistor is nearly equivalent to an ideal switch in series with a battery V_p and shunted by a current source I_p. This battery and current source impose a practical lower limitation on the power level of a load circuit which can be satisfactorily controlled by a transistor switch. The closer that this point P can be made to the origin, the lower the satisfactory operating level becomes. It was first found experimentally and later verified theoretically that, if the collector and emitter terminals were interchanged, the resulting characteristics had an intersection much closer to the origin than the normal connection. This connection has been termed the inverted connection and will be analyzed in detail. Since the effective alpha for the connection is typically about 0.5, its gain is not as high as that of the normal connection, but in many low-level circuits this is not of significant consequence.

The low-level operation of a fused junction transistor can be analyzed quite accurately by using the set of equations which were first given by Ebers and Moll.[1] These equations do not take into account any internal resistance drops, in that the potentials that appear are not the applied potentials but are the potentials across the junctions. For the moment, these internal resistance drops will be neglected and the potential across the junction is assumed to be the same as the applied potential.

Before presenting these general transistor equations, it is worth while to discuss the theoretical equation for the behavior of a fused junction diode which is

$$I = I_s\left(\epsilon^{\frac{qV}{kT}}-1\right)$$

where

q = charge of electron
k = Boltzmann's constant
T = absolute temperature, K
$\frac{kT}{q}$ = 26 microvolts at 20 C
I_s = constant

If V is made more than 1/10th volt negative, the exponential term becomes negligible with respect to unity and the leakage current theoretically remains constant at I_s up to the breakdown voltage. In practice, the leakage increases somewhat with voltage but many diodes approximate this theoretical behavior quite closely. This constant leakage current I_s is called the reverse saturation current of the diode; this has nothing to do with the conduction saturation previously discussed. In the forward direction, the current is nearly an exponential of the applied voltage. Note that the ratio of the forward-to-back currents at any particular voltage $\pm V$ is given by

$$\frac{I+}{I-} = \frac{\epsilon^{\frac{qV}{kT}}-1}{\epsilon^{-\frac{qV}{kT}}-1} = 55 \text{ at } \pm 0.1 \text{ volt}$$

and is independent of any of the properties of the semiconductor or the geometry and size of the unit. In other words, a unit having a low forward drop will have a high leakage and vice versa.

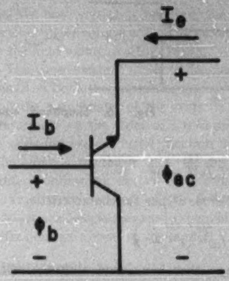


Fig. 14. Equations for the inverted connection

$$I_b = +I_{c0}\left(\epsilon^{\frac{q\phi_b}{kT}}-1\right)-(1-\alpha_n)I_e$$

$$I_e = I_{0e}^{\frac{+q\phi_b}{kT}}\left(\epsilon^{-\frac{q}{kT}\left(\phi_{ec}-\frac{kT}{q}\ln\alpha_n\right)}-1\right) = \left(\frac{1-\alpha_n}{\alpha_n}\right)I_0$$

where

$$I_0 = \frac{\alpha_n I_{c0}}{1-\alpha_i\alpha_n} = \frac{\alpha_i I_{c0}}{1-\alpha_i\alpha_n}$$

The corresponding theoretical equations for a transistor are given in Fig. 13. From the construction of a transistor it might be expected that it would behave as two diodes which react on one another and the similarity of the terms in these equations to that just given for a diode confirms this viewpoint.

As was mentioned, the most favorable connection for low-level operation is the inverted circuit shown in Fig. 14. The equations governing this connection can be derived from those previously given by using the relations: $\phi_b = \phi_c$, and $I_b = I_e - I_c$. With appropriate juggling, these expressions can be put in the forms given in Fig. 14.

By recognizing that an ideal diode has the characteristics $I = I_s[\epsilon^{(q\phi/kT)}-1]$, these equations may be interpreted as being equivalent to the characteristics of the circuit shown in Fig. 15, where diode D_1 has the characteristic

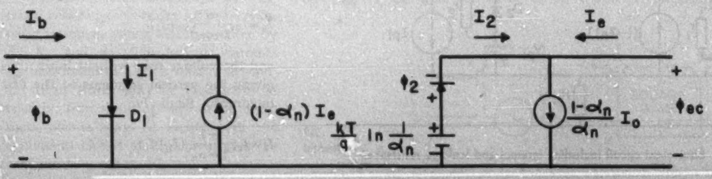


Fig. 15. Large signal equivalent circuit

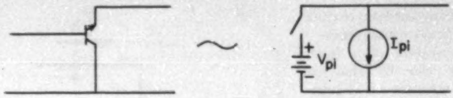


Fig. 16. Simplified equivalent circuit

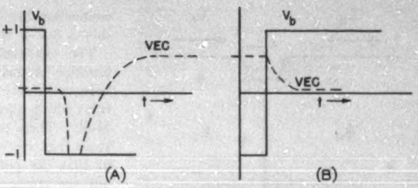


Fig. 18. Transient phenomena

$$I_1 = I_{c0}\left(\epsilon^{\frac{q\phi_b}{kT}} - 1\right)$$

and diode D_2 has the characteristic

$$I_2 = I_{e0}{}^{\frac{q\phi_b}{kT}}\left(\epsilon^{\frac{q\phi_2}{kT}} - 1\right)$$

The author has found this equivalent circuit extremely valuable in visualizing the behavior of switching transistors.

For example, it can be shown that the values of the battery and current source in the load circuit are equal to the co-ordinates of the common intersection point $P(I_{pi}, V_{pi})$ of the emitter-collector volt-ampere characteristics. To obtain an expression for the corresponding co-ordinates for the common emitter connection, it is only necessary to interchange the subscript i with the subscript n and I_{e0} with I_{c0}. Note that since I_0 is symmetrical in these terms, it remains invariant and the co-ordinates of P for the two connections become:

Inverted connection

$$\frac{1-\alpha_n}{\alpha_n}I_0, \ \frac{kT}{q}\ln\frac{1}{\alpha_n} = (I_{pi}, V_{pi})$$

Common emitter connection

$$\frac{1-\alpha_i}{\alpha_i}I_0, \ \frac{kT}{q}\ln\frac{1}{\alpha_i} = (I_{pn}, V_{pn})$$

For most commercial junction transistors, α_n is much nearer unity than α_i, typical values being $\alpha_n = 0.99$; $\alpha_i = 0.5$.

Equating the connection expressions for these values:

Inverted connection

$$\left(0.01\,I_0,\ 0.01\frac{kT}{q}\right)$$

Common emitter connection

$$\left(1.0\,I_0,\ 0.69\frac{kT}{q}\right)$$

This illustrates that the co-ordinates of the intersection P for the inverted connection may be orders of magnitude smaller than those for the common emitter. This fact has two very important consequences:

1. A transistor, when used in the inverted connection, will function at much lower load power levels than when used in the normal connection.
2. A transistor, when used in the inverted connection, has a room-temperature leakage current at cutoff which is an order of magnitude lower than that of the normal connection.

This second property has the important consequence that the value of $AVI_{20}(V)$ at any given voltage may be an order of magnitude less for the inverted than for the normal connection and, therefore, the inverted connection may be operated at a much higher temperature without runaway.

The switching function is closely associated with the properties of the diode D_2 which has the very interesting characteristic $I_2 = I_0 \exp(q\phi_b/kT)[\exp(q\phi_2/kt) - 1]$. Consequently, it behaves like an ideal diode. Its unique property is that the coefficient $I_s = I_0 \exp(q\phi_b/kT)$ can be varied by means of another electric potential ϕ_b. Thus, this coefficient can be made large during conduction and very small during cutoff. It is quite feasible for this diode to present a ration of front-to-back conductance on the order of 1,000,000 down to a few microvolts. In practice, the reduction of back leakage is limited by the voltage-sensitive component of the emitter leakage which can be represented by a resistance in parallel with the diode and which is ordinarily on the order of 1 megohm. In most low level applications, the base driving voltage is large compared to the current to be controlled and this diode D_2 can be thought of as an ideal switch controlled by the voltage ϕ_b.

Thus, if the base current is sufficient to hold the transistor saturated during the conduction period, a junction transistor having a square wave driving voltage impressed between the base and collector (the so-called inverted connection) is approximately equivalent to a switch in series with a battery and in parallel with a current source, as shown in Fig. 16.

Typical magnitudes are about 1 millivolt for V_{pi} and 1 microampere for I_{pi} at room temperature. V_{pi} varies somewhat with temperature but generally remains between 1/2 and 2 millivolts over the range of −65 to 100 C. I_{pi} increases exponentially with temperature at about 8 per cent per C.

The normal connection in which the square wave is applied between the base and emitter has a similar equivalent except that V_{pn} and I_{pn} are usually one to two orders of magnitude greater than V_{pi} and I_{pi}.

An analytical expression for the potential drop at any current is given by

$$\phi_{ec} = V_{pi} + \frac{kT}{q}\ln\frac{1+(1-\alpha_n)\dfrac{I_e - I_{pi}}{I_b + I_1}}{1-\dfrac{1-\alpha_i}{\alpha_i}\dfrac{I_c - I_{pi}}{I_c + I_1}}$$

where

$$I_1 = I_0\left(\frac{1-\alpha_i}{\alpha_i} + \frac{1-\alpha_n}{\alpha_n}\right)$$

This is, in fact, the equation for the entire family of curves of Fig. 12. When the denominator of the logarithm becomes zero

$$I_e = I_{pi}\frac{\alpha_i}{1-\alpha_i}(I_b + I_1)$$

and the voltage rises without limit, thus giving the vertical segments in the first quadrant. Similarly

$$I_e = I_{pi}\frac{-1}{1-\alpha_n}(I_b + I_1)$$

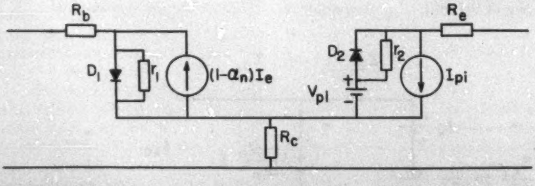


Fig. 17. Equivalent circuit including internal and leakage resistances

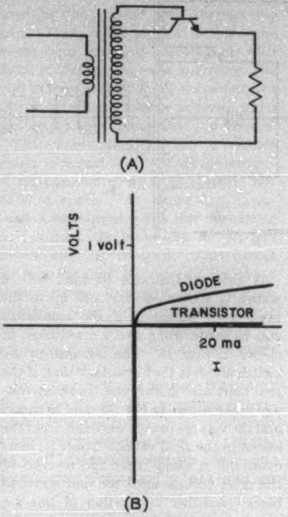


Fig. 19. Synchronous rectifier

corresponds to the vertical portions in the third quadrant. If the maximum value of $I_e - I_{pt}$ in the first quadrant is termed the maximum forward current and is denoted by the symbol $I_f = \alpha_i/(1-\alpha_i)$ $(I_b + I_1)$ and the maximum reverse current by $I_r = 1/(1-\alpha_n)$ $(I_b + I_1)$, then the expression for ϕ becomes simply

$$\phi_{ec} = V_{pt} + \frac{kT}{q} \ln \left(\frac{1 + \frac{I_e - I_{pt}}{I_f}}{1 - \frac{I_e - I_{pt}}{I_r}} \right)$$

The dynamic resistance $R_{dyn} = \phi_{ec}/I_e$ evaluated at the point $P(V_{pt}, I_{pt})$ is $(kT)/[q(I_b + I_1)]$ $(1-\alpha_i\alpha_n)/\alpha_i$ and this resistance may be added in series with the circuit of Fig. 16 to give a more accurate equivalent. Ordinarily, I_1 is completely negligible compared to I_b during the conduction period and hence, may be omitted from these equations. If $(1-\alpha_i\alpha_n)/\alpha_i = 1$ and $I_b = 10$ ma, then $R_{dyn} = 2.5$.

In practice, there are always some internal resistances in series with the transistor terminals which will be denoted by R_b, R_c, and R_e. The leakage currents are also found to have a voltage-dependent term which may be thought of as resulting from shunting resistances. Both these effects may be included in the equivalent circuit of Fig. 15 to give the configuration of Fig. 17.

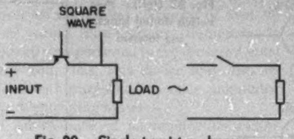


Fig. 20. Single-transistor chopper

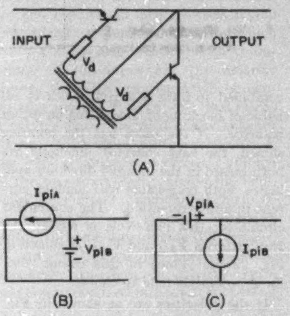


Fig. 21. Two-transistor chopper

R_b is usually quite appreciable, 50 to 200 ohms, and may vary by a factor of 4 from one unit to another of the same type. R_c and R_e are on the order of 1 ohm for fused junction units but R_c may be hundreds of ohms for grown junction types. Since R_e serves to couple the input and output circuits, grown junction units are generally not applicable in low-level circuits. R_1 and R_2 may be on the order of 0.1 to 10 megohms.

The transient behavior of junction transistors has been beautifully analyzed by J. L. Moll[3] and this work will not be repeated here. The usual manifestations of the transient phenomena in a low-level common emitter circuit are: If the unit is initially conducting with a very small voltage drop V_1 and the base potential is abruptly reversed, the collector emitter potential will remain unchanged for a short interval and will then suddenly drop to approximately the value of the applied base potential and then decay to its steady value; see Fig. 18(A). Similarly, if the unit has been cut off and the base voltage is reversed, the collector-emitter voltage will decay smoothly to the final value; see Fig. 18(B).

The time constant of these decays depends on both the transistor parameters and the magnitudes of the currents and are usually on the order of 1 to 5 microseconds for present fused junction units. The effect, depicted in Fig. 14, may be quite troublesome in low-level applications although several of the following circuits are designed so that these voltage spikes do not appear in the output.

Applications

A circuit which very nicely demonstrates the advantages to be gained through the use of transistors is the synchronous rectifier shown in Fig. 19. When the instantaneous polarities are as shown, the transistor is conducting in a common collector connection and can thus handle a large current and yet have the voltage drop across it very small. When the polarity reverses, the transistor blocks and only the small leakage current I_{pt} flows. This device thus avoids the severe inner dependence of the forward and back characteristics of a diode and can exhibit forward-to-back ratios of 100,000 down to 1 millivolt. As a consequence of the reduced leakage, this rectifier can operate at temperatures about 40 C hotter than a corresponding diode of the same material.

If this circuit is used as the self-saturating rectifier in a magnetic amplifier, the voltage that the load windings is required to sustain may be reduced to a fraction of 1 volt, thereby making practical the use of very small cores to yield an extremely sensitive amplifier having a high speed of response. The low-voltage characteristics of the transistor make possible the replacement of the input mechanical vibrators in low-level d-c amplifiers. One of several possible circuits is shown in Fig. 20.

If the transistor were an ideal switch, this circuit would be equivalent to the right-hand figure. Here, the input is alternately connected and disconnected to the output and, therefore, the output voltage is a rectangular wave that alternates between zero and the value of the input voltage. Actually, the transistor is not

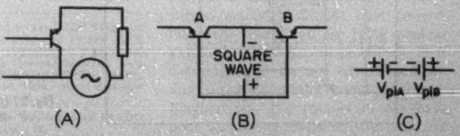


Fig. 22. Two-transistor a-c switch

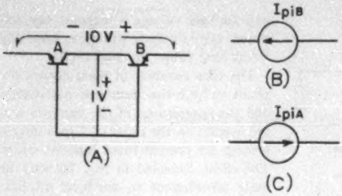


Fig. 23 (left). A-c switch during blocking interval

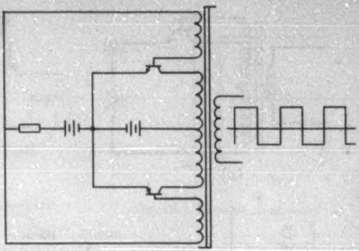


Fig. 26 (right). D-c to a-c converter

ideal, but is roughly equivalent to a battery: V_{pi} when conducting and a current source, I_{pi} when nonconducting. Thus, even if the input voltage is zero, a voltage of V_{pi} will appear at the output during the conducting half-cycle, and a voltage $I_{pi}R$ during cutoff. If this latter voltage is to be kept small, it means that the load resistance R must be small.

A much better modulator circuit is given in Fig. 21(A). First consider the behavior of this circuit with the input short-circuited. The two half-cycles can then be represented by the equivalent circuits of Figs. 21(B) and (C). The current sources are effectively short-circuited by the batteries during both halves of the cycle and the output alternates between V_{piA} and V_{piB}. If these two voltages are equal, the output is pure direct-current and contains no a-c component. Thus, as far as inverter operation is concerned, V_{piA} and V_{piB} tend to cancel.

An NPN transistor can block only if the base is more negative than both the collector and the emitter. Therefore, a transistor in the simple circuit described in the following cannot block an alternating source during the half-cycle when the polarity is as shown in Fig. 22(A). Fig. 22(B) shows how two transistors can be combined to form a switch which is capable of blocking or conducting on either polarity. If the wave has the polarity shown, the base collector junctions are both biased in the forward direction and, hence, both transistors will conduct current in either direction. The equivalent circuit for this half-cycle is merely the two batteries V_{piA} and V_{piB}, as shown in Fig. 22(C). Thus, in this circuit also, these potentials tend to cancel.

If the polarities are as shown in Fig. 23(A), both collector junctions are biased in the reverse direction. The emitter of transistor A is also negative with respect to the base and, hence, both junctions of A are biased backwards and the circuit is cut off. Similarly, if the controlled circuit potential reverses, transistor B has both junctions biased backwards and, hence, blocks the circuit. The equivalent circuit for the cutoff conditions is quite complicated, but if the controlled circuit potential is 1 volt or more and has the polarity shown, the equivalent circuit is approximately given by Fig. 23(B) and, similarly, for the opposite polarity by Fig. 23(C).

A chopper that can be used with a source of either polarity and up to the breakdown potential of the transistors can be made using two a-c switches, as shown in Fig. 24. The 2-transistor a-c switch also has the advantage that, if the two units are similar, the transient-voltage peaks shown in Fig. 24 tend to cancel and the voltage decays smoothly from the initial to the final value. There is, however, still a decay after the reversal of the drive voltage until the unit starts to block. Another application of this a-c switch is in the phase-sensitive detector or demodulator, as shown in Fig. 25.

The driving square wave establishes the phase reference. If the square wave polarity is as indicated, transistor switch B, will be conducting and transistor switch A will be cut off. Thus, the circuit will be completed, as shown by the dashed line; the direction of the current will depend upon the polarity of the source. On the other half-cycle, A will conduct and B will be cut off, thereby

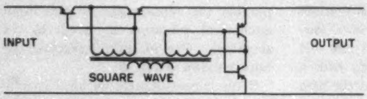


Fig. 24 (left). Four-transistor chopper

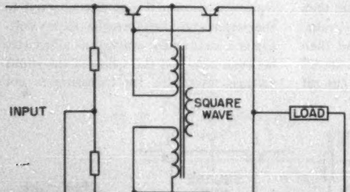


Fig. 25 (left). Four-transistor demodulator

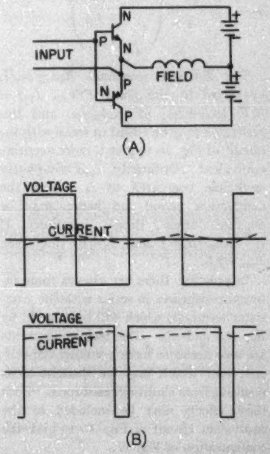


Fig. 27 (right). Transistorized field controller reversing the current path through the load. Thus, if the input is a sine wave whose zero crossings correspond with those of the square wave reference, the load voltage will be a full wave rectified sine wave whose polarity reverses when the input phase changes by 180 degrees.

This circuit requires a low impedance source to assure that the leakage current of the nonconducting transistor will not cause an appreciable voltage drop to appear at the output. If a low impedance load is used, the input and output terminals may be reversed. This circuit can also be used to measure the phase angle between the input and reference voltages. This demodulator is very linear from about 100 microvolts to 50 volts and is operable over a temperature range of −65 to +75 C.

By combining switching transistors and elements having rectangular hysteresis core materials, it is possible to achieve many novel results. For example, Fig. 26 shows a converter which produces a rectangular a-c output with an over-all conversion efficiency of over 90 per cent. The output frequency is directly proportional to the d-c input voltage and, thus, this device may also be used as a precision frequency modulator in telemetering systems.

An example of an application motor control system is the static field controller shown in Fig. 27(A). Here, advantage is taken of the fact that a transistor can conduct in either direction and can, therefore, perform functions which are not at present practical with thyratrons. If the input has the polarity shown, the NPN transistor will be conducting and the PNP cut off; this places a voltage across the field of the polarity shown. With the opposite polarity input, the PNP conducts and places a voltage of the opposite polarity across the field. Thus, this circuit functions similarly to a double-pole double-throw reversing switch with one outstanding difference—it can be thrown practically instantaneously. In effect, the polarity of the field can be reversed without ever opening the field circuit; thus, there will be no high-voltage surges and no discontinuities in the current. The average field current can be controlled by varying the ratio of the time that one polarity of voltage is applied to the time that the polarity is applied. Typical wave forms are shown in Fig. 27(B). If the frequency of these rectangular waves is high compared to the reciprocal of the time constant of the field, the field current will have very little ripple. This circuit can be held closed in either direction and can apply full forcing to the field in either direction. These applications serve to illustrate the great potentialities of using fused junction transistors as switches at both high- and very-low-power levels.

References

1. Large-Signal Behavior of Junction Transistors, J. J. Ebers, J. L. Moll. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 42, no. 12, Dec. 1954, pp. 1761–72.

2. Large Signal Transient Response of Junction Transistors, J. L. Moll. *Ibid.*, pp. 1773–84.

Transistor Demodulator for High-Performance Magnetic Amplifiers in A-C Servo Applications

R. O. DECKER
ASSOCIATE MEMBER AIEE

FULL-WAVE magnetic amplifiers are not efficiently controlled with a-c signals applied directly to the control windings because the circuit configurations necessary do not give sufficiently high gain. Half-wave magnetic amplifiers which can accept a-c signals directly have relatively low gain. The most sensitive, stable magnetic amplifiers require a d-c control current, but in many servo applications the servo amplifier must accept a phase-reversible alternating voltage and deliver an output that will reverse phase as the phase of the input signal reverses.

An ideal solution to the a-c input problem is obtained by inserting a phase-sensitive demodulator in series with the magnetic amplifier control windings. Such a demodulator however, must have good balance characteristics over a wide temperature range, and its transfer curve must be linear even at low voltage levels. To maintain the reliability of the amplifiers, the use of vacuum tubes should be avoided. Demodulators using selenium rectifiers in a balanced bridge arrangement are available commercially, but the unbalance with zero input signal is generally intolerable for low-level servo applications. The relatively high threshold voltage of a selenium rectifier makes the demodulation of signals below several hundred millivolts difficult and good linearity is difficult to obtain at low voltage levels.

A demodulator circuit that does not suffer from the disadvantages of the selenium rectifier units but that avoids the use of vacuum tubes is shown in Fig. 1. It is a transistor demodulator that has been combined with a stable, high-gain 2-stage magnetic amplifier to produce a unit that is very effective in a synchro-controlled position system.

Operation

The demodulator circuit is one of a newly developed class of transistor circuits in which the transistors are used as switches.[1] An alternating reference voltage is applied between the base and collector of the PNP transistors. When the base is positive with respect to both the emitter and collector, the transistors will act as open switches with very low leakage. When the base is negative with respect to either, the transistors will act as closed switches with very low voltage drop.

The modes of operation are illustrated in Fig. 2. The transistor pair, when acting as a closed switch, will pass current in either direction if the reference source supplies sufficient base current. Since the transistors are being used as switches rather than as amplifiers, their characteristics are not critical and the units do not have to be carefully matched. Furthermore, the transistors can have poor amplifying characteristics and still

---

Paper 55-67, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing December 10, 1954.

R. O. Decker is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

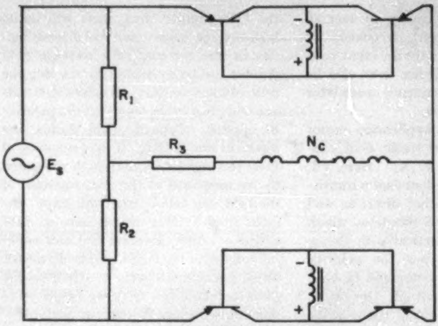


Fig. 1. Circuit diagram of transistor demodulator

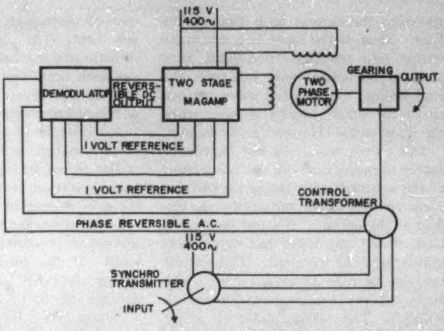


Fig. 3. Position servomechanism using transistor demodulator combined with 2-stage magnetic amplifier

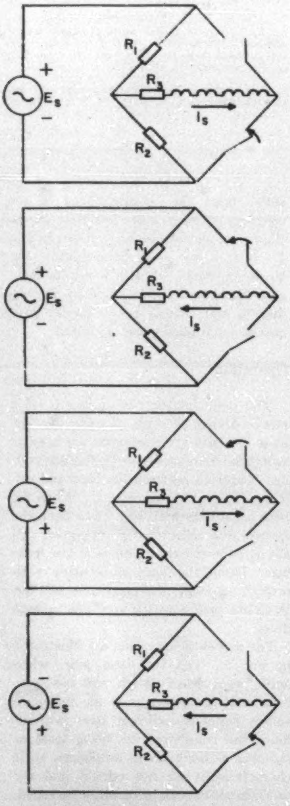


Fig. 2. Modes of operation of transistor demodulator

be successfully used. The property which is most important is the Zener voltage characteristics of the transistors. The maximum signal voltage E_s appears across the emitter-collector junction of one transistor during the time in which one pair of transistors is acting as an open switch; thus, the junction must be able to withstand the peak voltage which the signal source can supply. For many synchros this maximum voltage is 30 volts rms, and the transistors must then have a Zener voltage of 42 volts.

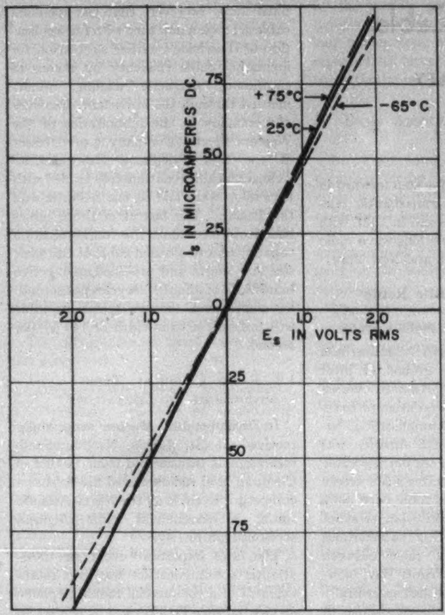
In the circuit of Fig. 1, two resistors, R_1 and R_2, are shown; these are used to provide a split signal source. However, in some applications, these can be replaced by a split secondary of an input transformer. The two resistors R_1 and R_2 do not have to be carefully matched because, as can be seen from Fig. 2, for either phase of the alternating voltage to be demodulated, the circuit operation must be symmetrical and the balance of the unit will not be affected by an unbalance in R_1 or R_2. The nominal value for this application is 10,000 ohms because the recommended minimum load for a synchro control transformer is 10,000 ohms. It can be seen from Fig. 2 that, for very low values of load resistance, the minimum resistance seen by the signal source is $R_1 = R_2$. For other applications, the resistors R_1 and R_2 may be chosen to be any value with the minimum value being set by either the loading of the signal source or the maximum permissible current of the transistors. Of course, the resistance of the load will have to be considered also because of the current division between the load and bridge resistances. R_3 in many applications is the resistance of the control windings, with no external resistance being necessary.

The reference voltage needed for switching the transistors is approximately 1 volt and is supplied from two separate secondary windings of transformer T_1. The reference voltage should be as nearly as possible in phase with the voltage of the signal source because the transistors block only when the proper polarities exist between the various terminals. In the experimental unit discussed subsequently, the reference voltages were obtained by adding two additional secondary windings to the supply transformer of the first stage of the 2-stage magnetic amplifier.

Experimental Characteristics

To illustrate a practical application, the demodulator has been combined with a 2-stage magnetic amplifier that has a high-volt-per-ampere-turn gain when direct current is supplied to the control windings. A typical system is shown in Fig. 3. The control transformer will supply a phase-reversible 400-cycle voltage that will be a measure of the relative position of the shafts of the control transformer and synchro transmitter. The voltage gradient of the control transformer is 400 millivolts per degree with a maximum output voltage of 27 volts.

The d-c output current versus applied voltage characteristic of the demodulator was measured at room temperature, −65 degrees centigrade and at +75 degrees centigrade. The unit was kept at the extremes of temperature for 4 hours before the measurements were made. The curves are shown in Fig. 4. With R_3 equal to 2,000 ohms, the null point shift under temperature extremes is less than 1



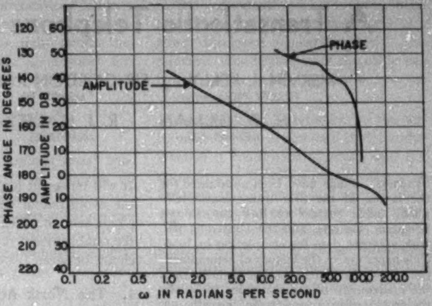


Fig. 4 (left). D-c output current versus a-c signal voltage characteristics of transistor demodulator

Fig. 5 (above). Frequency - response characteristics of position servomechanism

microampere. The data plotted in Fig. 4 were taken with voltage applied from the control transformer. In this particular system, control transformer voltage is not exactly in phase with the reference voltage applied between the transistor terminals; thus, the rectification efficiency of the demodulator is 78 per cent.

It will be noted that the characteristics remain linear even at small voltages. This linearity exists for voltages at least as small as 20 millivolts, which is the noise level of the synchro units. The influence of the magnetic amplifier on the demodulator, and vice versa, was tested by measuring the output current of the demodulator, both with the amplifier control winding short-circuited and with it connected. The difference was no greater than that caused by the resistance of the control winding alone. The effective resistance of the demodulator which determines the time constant of the first-stage amplifier is

$$\frac{R_1R_2}{R_1+R_2}+R_3$$

if the internal resistance of the signal source is neglected. For most applications, the effective resistance can be maintained at approximately 10,000 ohms, and the first-stage amplifier can be made to have a response time no greater than 3 or 4 cycles of the supply frequency. The first stage, which is usually a d-c to d-c amplifier, can then be treated as an operational magnetic amplifier.[2] Series lead or lag networks can be used with the d-c to d-c amplifier, or feedback through frequency sensitive networks around the amplifier can be used to provide system compensation.[3] Fig. 5 is a plot of the frequency-response characteristics of the system shown in Fig. 3. The amplifier response has been extended by negative integral feedback around the first stage and by a series compensating network between amplifier stages. The negative integral feedback network supplies a steady-state feedback to stabilize the amplifier against temperature, voltage, and frequency variations. By using negative feedback in this manner, it has been possible to reduce the drift level of a magnetic amplifier position servomechanism with a transistor demodulator to a position error of 0.1 degree over an ambient temperature range of −65 to +71 degrees centigrade, and with 10-per-cent variations in supply voltage and frequency.

Conclusions

The transistor demodulator described in this paper is a stable detector and demodulator capable of operating at signal voltage levels at least as low as 20 millivolts and as high as 30 volts rms with low drift over the temperature range from −65 to +75 degrees centigrade. It is extremely well suited for operation in a synchro-controlled system because it can be readily combined with a stable, full-wave magnetic amplifier.

The magnetic amplifier-transistor demodulator combination makes a stable, high-gain amplifier that will amplify a-c signals and still enable the full amplification capabilities of the magnetic amplifier to be utilized. For synchro-controlled systems, the sensitivity and drift stability are great enough so that errors caused by inaccuracy in the magnetic amplifier-transistor combinations are not limiting factors in the over-all servo performance.

References

1. The Use of Junction Transistors as Switches, R. L. Bright. *AIEE Transactions*, vol. 74, pt. I, Mar. 1955, pp. 111–21.

2. The 2nd Feedback Control Systems Conference. *AIEE Special Publication S-63*, Apr. 1954, "Operational Magnetic Amplifiers for Servomechanism Application," W. F. Horton, R. O. Decker, pp. 123–33.

3. Alteration of the Dynamic Response of Magnetic Amplifiers by Feedback Techniques, R. O. Decker. *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 658–65.

A Transatlantic Telephone Cable

MERVIN J. KELLY
FELLOW AIEE

SIR GORDON RADLEY
MEMBER IEE

G. W. GILMAN
MEMBER AIEE

R. J. HALSEY
MEMBER IEE

Synopsis: Plans have been announced for the laying of the first transatlantic telephone cable system to link the United Kingdom, Canada, and the United States, with a target date for completion in 1956. This paper traces the history of communication across the North Atlantic and discusses the inadequacy of radio circuits to satisfy the requirements of traffic growth. A general description is then given of the cable system which will provide 36 telephone circuits across the Atlantic and 60 between Newfoundland and Nova Scotia. The project has been made possible only by the development of submerged repeaters containing long-life vacuum tubes and other components, and these are described, as well as the American experimental work that led to the development of a flexible repeater housing which could be laid at depths of 2 miles without risk of damage to itself or to the cable. Fifty-two of these repeaters will be used in each of two 1-way cables, nearly 2,000 nautical miles long, connecting Newfoundland with Scotland. Newfoundland will be connected to Nova Scotia by a single cable containing 16 2-way repeaters of British design evolved from that which has been extensively used for schemes in home waters.

IN DECEMBER 1953 plans were announced by the American Telephone and Telegraph Company, the British Post Office, and the Canadian Overseas Telecommunication Corporation for the laying of the first transatlantic telephone cable system. The system will provide 36 high-grade telephone circuits between the United Kingdom, Canada, and the United States, and commercial operation is contemplated at the end of 1956.

In scope the project far exceeds anything that has been attempted hitherto; the two cables on the main crossing will each be nearly 2,000 nautical miles* long and each will contain 52 submerged repeaters, laid at depths of up to 2,000 fathoms, whereas at the time of writing the longest submarine telephone cable is less than 200 nautical miles long and no cable contains more than four repeaters. The American Telephone and Telegraph Company and the British Post Office have joint engineering responsibility for the project, which is the outcome of studies and experiments started many years ago. It will represent a major step forward in the evolution of the transatlantic telephone service which began in 1927 with the establishment of one long-wave radio circuit between London and New York.

1. The North Atlantic Route

1.1 The Early Telegraph Cables

Historically, the North Atlantic route has been the proving ground of most major advances in long-distance oversea communication. The first attempt to establish telegraphic communication between Europe and North America was made by cable in 1857–58, but the cable failed in October 1858 after a few weeks of service. In 1865 the route once more became the scene of activity, but it was not until the following year that the hazardous operation of laying a cable between Valentia, Ireland, and Trinity Bay, Newfoundland, was accomplished successfully and the North Atlantic route opened to electrical communication.

The successful completion of the first transatlantic telegraph cable was due to the foresight, courage, and determination of Cyrus W. Field, an American, coupled with an outstanding British technical accomplishment of its day. The communication capacity of the cable was, however, very limited. Dr. O. E. Buckley noted in the 1942 Kelvin Lecture[1] that, even with the improvements in terminal apparatus and operating techniques introduced by Sir William Thomson (later Lord Kelvin) and his associates, the first transatlantic cable operated at a speed of only 3 words per minute. This required a bandwidth of perhaps 1.5 cycles per second.

To meet the increasing demand for rapid communication across the North Atlantic which arose after the first cable went into operation, more cables were laid and there are now some 20 in operation on this route.[2] The best available scientific talent, particularly in the United Kingdom, was devoted to increasing the efficiency of the early cables, and late nineteenth century research on submarine telegraph cables and their terminal apparatus may well be one of the first examples of the large-scale effectiveness of applied research. It was not, until 1924, however, that the deep-sea cable art took a new turn with the application of Heaviside's earlier suggestion for increasing cable efficiency by means of continuous magnetic loading, implemented through the important discovery of permalloy in the laboratories of the Western Electric Company, a predecessor of the Bell Telephone Laboratories.[3] Telegraph cables continuously loaded with permalloy were laid in the Atlantic and the Pacific. The fastest of these cables had about four times the communication capacity of a nonloaded cable of the same size and length and provided an effective bandwidth of about 100 cycles per second. By comparison, the new telephone cable will transmit a bandwidth of 144 kc per second.

1.2 Radio Circuits: Their Advantages and Limitations

In December 1901 Marconi successfully received at St. John's, Newfoundland, radio signals transmitted from Poldhu in Cornwall, and radio entered the field as a competitor to cable on the North Atlantic route. A commercial radio-telegraph service began in 1908.

The next important event in transatlantic communication was the establishment of a commercial radio-telephone service between London and New York in January 1927, when the long-wave radio transmitting stations at Rugby and Rocky Point were brought into operation. The circuit used the single-side-band suppressed-carrier technique which has since found universal application for both radio and cable telephone circuits. Another important event was the opening of the short-wave (high-frequency) radio spectrum to long-distance communication. Short waves were introduced on the transtlantic route in 1928, and there are, at the time of writing, 14 short-wave radio

---

* One telegraph nautical mile or naut = 6,087 feet = 1.153 statute miles.

Paper 55-230, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing December 3, 1954.

Mervin J. Kelly and G. W. Gilman are with the Bell Telephone Laboratories, Inc., New York, N. Y., and Sir Gordon Radley and R. J. Halsey are with the Post Office of the United Kingdom, London, England.

The concept of a transatlantic telephone cable owes its origin to the inspiration of Dr. O. E. Buckley's early work, for he laid the firm technical foundation for such a project. The subsequent developments recorded in this paper have been the work of many engineers on both sides of the Atlantic; on the American side in the Bell Telephone Laboratories, the Western Electric Company, and the Simplex Wire and Cable Company; on the British side in the Post Office Research Laboratories, Standard Telephones and Cables Limited, and Submarine Cables, Limited. Without this work the particular project described would have been impossible.

telephone circuits between the United Kingdom and the United States and two between the United Kingdom and Canada. Both events were to have an indirect but important effect upon the subsequent development of submarine cables. The availability in the early 1930's of comparatively wide-frequency bands for short-wave radio communication, the economic depression of that period, and other factors combined to defer whatever plans there may have been to lay additional telegraph cables across the North Atlantic. On the other hand, recognition of the technical limitations of radio for transatlantic telephony soon led to studies of the feasibility of a North Atlantic submarine telephone cable.

The advantages of short-wave operation have been:

1. Comparatively low-cost installations.
2. Flexibility, i.e., ease of establishing taking down, and rerouting circuits.
3. Direct access country to country.

With respect to the last point, it is of interest to note that arrangements have come into effect to route certain transatlantic traffic through radio-relay stations favorably located to avoid zones where ionospheric conditions are most frequently disturbed. On the other hand, the following disadvantages of short waves are familiar:

1. Although, when ionospheric conditions are stable, short-wave operation furnishes satisfactory transmission, no way has been found to provide day-to-day continuity and reliability comparable to that of good wire lines.

2. The deficiencies of short-wave operation are particularly marked on the North Atlantic route, which is exposed to severe ionospheric disturbances. (The past 25 years have been characterized by three periods of very severe disturbance. The first period covered about 3 years around 1930; the second, more moderate in effect and of shorter duration, centered at 1944; and the last, which is particularly serious started in 1950 and is still in progress.)

3. The short-wave radio spectrum has become overpopulated with time. Interference between services is an increasingly serious problem and, even when full allowance is made for the development of multichannel working on a single transmitter the number of transatlantic circuits could not be increased very greatly.

The original long-wave radio-telephone circuit, which is still in use, renders an important service in providing limited communication facilities when the short-wave circuits are not serviceable. However, the entire long-wave spectrum has a frequency bandwidth of less than 100 kc per second and, since this must be shared by the world's communication services (telephone and telegraph) which need long waves, long-wave radio cannot be regarded as more than an adjunct for transatlantic telephony.

2. Early American Work

In 1919 a small American group was organized under Dr. O. E. Buckley to study ways of increasing the transmission capacity of long, deep-water submarine cables; this was in accordance with the emphasis in all American submarine cable work on the problems of deep-water cable technology.

The first work of this group was, however, with short lengths of telephone cables at medium depths. In 1921 three iron-wire (Krarup)-loaded cables were laid between Havana and Key West, a distance of just over 100 nautical miles. These were the first cables made with an outer copper return conductor giving a coaxial structure. In deep-sea telegraph cable practice the custom had been to depend upon the sea itself and the armor wires of the cables as a return medium, but this is not satisfactory at telephone frequencies. Each of the three Key West-Havana cables provided sufficient frequency bandwidth to accommodate one telephone and three carrier telegraph circuits in addition to a d-c telegraph circuit. Two years later two cables were put into operation over the 23-nautical-mile route between the California shore and Catalina Island, and these were ultimately operated at frequencies up to 35 kc per second. The Key West-Havana and the Catalina cables were operated as 2-way cables, i.e., each cable was used for both directions of transmission simultaneously. Separation of the two directions of transmission was by means of balancing circuits at the terminals. There soon followed the developments in long, deep-water, high-speed submarine telegraph cables which were mentioned earlier.

By the late 1920's developments in submarine cable technology and in magnetic and insulating materials reached a point where a single-circuit telephone cable across the North Atlantic between Ireland and Newfoundland (a distance of 1,800 nautical miles) seemed feasible. A specific proposal to this effect was formulated by the Bell Telephone System in 1928 and discussed with the British Post Office. The proposal called for laying a single, continuously loaded, nonrepeatered coaxial cable, operating at a top frequency of about 3,800 cycles per second and providing a single telephone circuit. The central conductor was to have been loaded with wrappings of tape made of a newly developed magnetic material, perminvar. The insulation of the cable was to have been paragutta, also a new development.

As a check on the estimates of the performance of this design, and to determine whether it could be successfully laid and recovered, a 20-mile section was manufactured in Germany and laid in deep water in the Bay of Biscay in 1930. This was subsequently recovered and relaid in shallow water. A length of nonloaded cable with paragutta insulation was laid the same year between Key West and Havana. The projected cable was never laid across the Atlantic, principally because of economic conditions. However studies and experiments in the laboratory and at sea, looking toward a new and better cable, continued during the depression following 1930.

During the laying and recovering operations in the Bay of Biscay in 1930, "kinking" of the cable and buckling of the conductors occurred on several occasions. This extreme local deformation was caused by unwrapping of the armor wires at points along the cable. The damage is caused in the following way. As the cable passes aft of the ship, variations in tension cause twisting and untwisting of the cable structure; this twisting is harmless if it is distributed along the length of the cable, but if it is localized for any reason buckling can occur. Stopping and starting of the cable ship, or the existence of any large irregularity along the cable itself, are potential causes of damage; e.g., severe local distortion was produced artificially during the Bay of Biscay tests when a large board, nicknamed the "barn door," was lashed to the cable.

The structure of the proposed coaxial transatlantic telephone cable was inherently much more complex than that of the conventional nonloaded deep-water telegraph cable. The cable was therefore more subject to damage while being handled at sea and, because of the serious consequences of such damage, great attention was paid to the problem of maintaining the structure of the cable unaltered during laying operations. Laboratory studies led to a design in which the lays of the tape loading, the tapes of the copper inner and outer conductors, and of the armor wires were in the same direction. This was a departure from established thinking at the time, but has been carried through in the subsequent cable designs of the Bell Telphone Laboratories.

3. Multichannel Transatlantic Cable Project: Background of American Experience

By 1928, when the proposal for a transatlantic telephone cable was first formulated, electronic techniques were well established in land cable practice, although they had not reached the point when serious proposals could be made to lay a transatlantic submarine cable with submerged repeaters lying on the ocean bottom. For this to become feasible further improvements in technology were required, notably the use of amplifiers stabilized by negative feedback and pentode tubes with equipotential and uncombined cathodes. These developments, which were to make long-distance multicircuit carrier telephony a reality within the next 10 years, as well as still greater reliability, were needed to make the transatlantic plan practicable. Until these improvements became available, the only kind of repeatered transatlantic system that could be imagined was one in which the repeaters would be mounted on moored platforms at sea or would be in submerged buoys, fed with power from local batteries and visited at intervals for maintenance and repair. Systems of this kind were considered and discarded as impracticable.

By 1932 great advances were being made, and the first experiments on long-distance broad-band systems for use on land were well under way. While the solution of many of the submarine cable problems was not yet clear, the rapidly developing electronic art created the atmosphere of confidence that made it possible to contemplate seriously the design of a long-distance deep-water cable associated with submerged repeaters in which were long-life tubes which would be capable of transmitting a number of carrier telephone channels instead of only one channel.

The broad concept which grew up concerning the transatlantic cable system was the use of two nonloaded coaxial cables, one for each direction of transmission, into which repeaters would be spliced at regular intervals. Detailed discussion of the merits of two cables versus one is outside the scope of this paper. Suffice to say that an economic field for the use of twin cables (1-way operation in each) may be on long deep-water routes with traffic growing rapidly beyond the capacity of a single-cable system. In this case, however, the choice was influenced by the need to use repeater housings which would cause the minimum irregularity in the cable structure. As will be pointed out in section 4, 2-way repeaters for use in a single-cable system require more space in the repeater housing than do repeaters for 1-way operation.

The repeaters required for the transatlantic cable system had to be designed to withstand the shocks of laying and recovery and also the pressure of water encountered at the greatest depth on the North Atlantic route. The tubes used had to operate at sufficiently low cathode power and anode potential to make it feasible to supply power to the repeaters over the cable from the shore ends at a safe working potential. This potential was eventually established at roughly 2,000 volts d-c. Finally, the repeaters had to be long-lived enough to operate with small likelihood of failure over a period of time (generally taken as 20 years) which would be long enough to make the enterprise as a whole technically sound and economically attractive.

3.1 Repeater Housing

Early experience of the potential hazards of damage to, or modification of, the structure of the cable during handling led to the objective of a flexible housing, structurally as nearly as possible like the cable, and which could, without damage to itself or the cable, be passed over the drum and sheaves of the cable ship and paid out without slowing up or stopping the vessel. The so-called "snake-type" housing for the tubes and other repeater elements was developed; it will be described in section 7.2.

3.2 Long-Life Tubes

The tube finally adopted was a conservatively designed suppressor-grid pentode whose details were finalized in 1941 and which has not since been changed in any essential way. It is a tube with large cathode area, low cathode temperature (630 degrees centigrade), low plate potential (60 volts), large internal spacing between elements, and of generally rugged and shockproof construction. Some of these tubes have been under continuous test for a period of over 13 years and 18 have been in operation for 4 years in repeaters on the most recent Key West-Havana cables.[4] All indications are favorable for something better than 20-year life expectancy for North Atlantic service.

3.3 Components

In addition to the tubes, about 60 other circuit elements have to be accommodated inside the submerged repeater housing. These include resistors, capacitors, inductors, transformers, and crystals. Like the tubes, these elements must be designed and fabricated so as to avoid all possible risk of failure in service. They must also meet requirements of ruggedness and reasonably small size. Cost, fortunately, is a less important factor than it is in most other applications.

The failure of any one of the other components could be as serious as that of a tube, and a 5-step philosophy was evolved in the Bell Telephone Laboratories for their selection and manufacture, aimed at achieving as near perfection as humanly possible. The steps are as follows:

1. Draw on the background of experience with similar parts in use in the telephone plant.
2. Supplement experience with extraordinary care in selecting material and with special testing of components during production and after completion.
3. Use an open and simple structure to avoid hidden errors.
4. Select and train personnel for perfection of skill and pride in product.
5. Develop the worker's faculty of self-criticism so that only those parts are submitted which truly represent his best efforts.

In this way there was incorporated into the development of each component all the knowledge and integrity that American telephone practice and experience provided. Wire-wound resistors, for example, have been made for 50 years, and there are something like 50,000,000 of them in service. This type of resistor was therefore chosen for the submerged repeater in preference to the deposited-carbon resistor which is of more recent origin. Similarly, experience in the control of materials and quality production and experience in the telephone operating field were drawn upon in the selection of designs of permalloy-core inductors and paper and mica capacitors. This experience was supplemented wherever possible by accelerated life tests.

Care in the control of materials used for the repeater elements has been particularly important. A case in point is the selection of paper for the capacitors. Large quantities of this paper are purchased annually for the manufacture of capacitors for use in the general telephone plant; the paper is supplied in lots which are submitted to accelerated life tests. The lots which performed best in an interval of 6 months were made available to the Laboratories, where every roll was reinspected and only those showing the highest life were chosen for the capacitors.

3.4 Cable

The design of cable finally decided upon consists of a solid dielectric coaxial structure covered by the usual jute and armor for protection and strength. The composite central copper conductor consists of a central solid wire surrounded by a single layer of spiralling and abutting tapes carefully formed to fit closely about the wire with a minimum of voids. Such a composite structure has better flexibility and protection against breakage than a solid conductor, and has better a-c and d-c resistance than a stranded-wire conductor of the same outside diameter. The return conductor is also a flexible composite copper structure consisting of a single layer of abutting spiralling tapes, carefully formed into a tubular configuration. It is covered with an overlapping, spiralling copper tape for teredo protection. The various spiralling layers of tape and armor are all applied in the same direction of lay for reasons stated in section 2.

The original coaxial cable designs had used paragutta as the dielectric, but shortly before World War II representatives of the Post Office brought polyethylene (polythene), a British development of Imperial Chemical Industries, to the attention of the Bell Telephone Laboratories. On test it proved to have characteristics which made it superior to paragutta. It had a lower dielectric constant and higher dielectric strength, was more uniform in its characteristics, was less pervious to sea water, and had mechanical properties which made it more desirable. Polyethylene was soon adopted as a feature of the proposed transatlantic cable.

3.5 Key West-Havana Cables

During World War II the American group concerned with the cable was reduced to the barest minimum necessary to keep the development alive, but a machine to simulate cable-laying operations was constructed and used to test ideas concerning repeater housings.

Upon the cessation of hostilities the decision was taken to install on the Key West-Havana route a pair of cables of the design that had been evolved for transatlantic service. A special air-conditioned repeater assembly shop was set up at the Laboratories and a contract was entered into with a manufacturer for experimental cables. Arrangements were also made with the Western Union Telegraph Company for modification to the cable ship *Lord Kelvin*. By 1948 work had progressed to a point which enabled 20 nautical miles of cable and repeaters, together with special built-in testing facilities (accelerometers) for observing and measuring shocks to the cable during handling, to be laid in depths up to 2 nautical miles in the Bahamas. During this period of growing activity, methods of cable manufacture, including testing and quality control, were worked out in co-operation with the manufacturer.

This work culminated in May 1950 with the laying of the Key West-Havana cables (115 and 125 nautical miles).[4] The cable system includes six repeaters, five of which are on the ocean bottom at depths ranging from 20 fathoms to 950 fathoms. The sixth repeater is located in a cable vault on the shore near the Havana telephone office. These repeaters have been in operation, at the time of writing, for 4 years without failure and without indication of deterioration.

4. British Shallow-Water Experience

Many submarine telephone cables connect the United Kingdom to the continent of Europe. These are of a variety of types and lie in water which is nowhere deeper than 300 fathoms. The first coaxial cable of modern type, and generally similar to the cable described in section 3.4, was laid in 1937 and used paragutta dielectric. All postwar coaxial cables, however, have been constructed with polyethylene.

British development of submerged repeaters for existing cables has recently been described.[5–7] Experimental work by the Post Office did not start until 1938, but in 1943 a coaxial cable between Angelsey and the Isle of Man was equipped with a shallow-water repeater.[8] In 1946 a repeater was inserted into a cable between Lowestoft and Borkum to provide direct telephone communication to the British Armed Forces in Germany. These were the first uses of submerged repeaters for telephone traffic anywhere in the world. In these two early applications, GO and RETURN speech channels were provided over a single cable and separated on a frequency basis, the repeater being arranged to amplify only the channels transmitted in the high-frequency band.

The Post Office next turned its attention to the development of submerged repeaters suitable for operation in tandem and designed to transmit a supergroup of 60 circuits over existing coaxial cables.[9] At this stage a decision became necessary as to whether British development should continue to be in accordance with a pattern of transmitting GO and RETURN channels over a single cable, which was the established practice with nonrepeatered systems, or whether separate cables should be used for the two directions. Use of separate cables for the two directions of transmission meant greater simplicity of equipment, but Post Office experience with the inherent difficulties of submarine cable maintenance in the North Sea had led to a firm preference for making every cable self-contained in respect to circuits on which the traffic did not necessitate more than two or three cables on a route.

It was decided, therefore, that the repeaters should be designed for 2-way transmission. The primary supergroup of 60 speech channels is, by international agreement, assembled in the frequency band 312 to 552 kc per second, and the system was arranged to transmit this frequency band in one direction and the frequency band 24 to 264 kc per second in the other. High- and low-pass filters were necessary to separate GO and RETURN channels at each repeater, but these were arranged to feed into a common amplifier.

The directional filters, necessary for 2-way transmission, waste a part of the frequency spectrum, while the added components somewhat increase the probability of premature failure and require additional space in the housings. The loss of frequency spectrum and slightly increased risk of failure were not considered important in comparison with the operational advantages of single-cable systems, and the additional space required was not a serious problem because experience had shown that the large single-ended housings which were in use at the time could be laid and recovered without difficulty in shallow water.

The system evolved, which may include up to ten intermediate repeaters, appears adequate for all routes between the United Kingdom and Europe. It was first installed in 1950 and about 20 repeaters are in operation at the time of writing. Very similar 36-circuit systems have been in operation for 4 years between the Netherlands and Denmark.[7]

These earlier projects, while profitable and affording valuable experience, did not lead directly to a design suitable for a transatlantic project or for installation elsewhere at ocean depths. In 1948, however, the Post Office began to study the possibility of using submerged repeaters to increase the traffic capacity of the telegraph cables operated by Cable and Wireless Limited, and development of a housing suitable for installation in deep water was commenced in collaboration with a manufacturer. The design was modified after laying tests in 2,700 fathoms in the Bay of Biscay in 1951 which were carried through under ad-

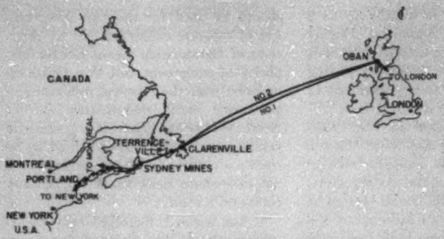


Fig. 1. Route across the Atlantic

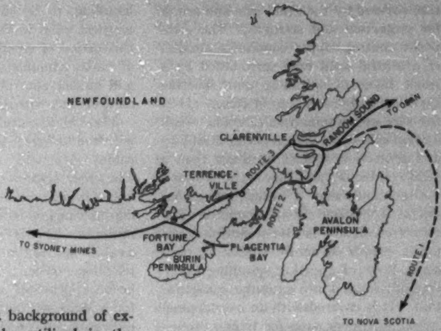


Fig. 2. Routes in Newfoundland

verse weather conditions and were not entirely successful. This housing, described in section 8.3, is rigid, like the shallow-water unit, but is double-ended and much smaller. It is in line with the cable, designed to transmit the full laying tension experienced in deep water, and is able to rotate freely during the laying operation. Limited tests were carried out in the Mediterranean Sea at depths of 250 fathoms in the autumn of 1953, when lengths of cable including the new repeaters were repeatedly laid and recovered. Twist, tilt, and recording devices included in one repeater housing established that the repeater could be handled and laid without significant bumping. It has yet to be established that the design is suitable for use in very deep water and that, in particular, the laying operation does not endanger the cable by the insertion of slack which subsequently tightens into kinks.

The Post Office did not commence serious study of the transatlantic problem until 1950 when it set up a committee to report on the future possibilities of repeatered telephone cables on this route. In 1952 it was decided to engineer a new telephone cable of about 300 nautical miles between Scotland and Scandinavia as a prototype for a system suitable for water of greater depth than that actually on the route, although this is no more than about 200 fathoms, and the requirements of 36 circuits on a single cable could have been met using the proved shallow-water repeater design. The cable was laid in the autumn of 1954. An innovation in submerged repeater circuit design is the use in each repeater of parallel connected amplifier units with two separate forward paths and a common feedback path. The tubes will be of a long-life type developed in the Post Office Research Laboratories for submerged repeater applications.

The decision to use deep-water-type repeaters for this system was an important step in building up a background of experience which will be utilized in the Newfoundland-Nova Scotia section of the transatlantic system.

5. Decision on Transatlantic Telephone Cable Project

Early in 1952 negotiations were once again opened between the American Telephone and Telegraph Company and the British Post Office. Two of the authors, Kelly and Radley, reported jointly on the status of submerged repeater development in the United States and in the United Kingdom after independent studies by engineers of the Company and of the Post Office.

At that time the American development had been completed in its basic features for several years. The tubes and other components, initially constructed or selected for reliability in service, had been under life observation since about 1940. Cable and repeater designs for a transatlantic project had been subject to stringent laboratory tests and had finally been given the practical trials, already referred to, on the Key West-Havana route. The authors concerned recommended that these designs should be used for the long link between Newfoundland and Scotland, because proved reliability is an essential requirement in a pioneering and costly venture of the difficulty of a transatlantic telephone cable. This is evident.

There had not been an opportunity to subject to comparable tests the elements in the British development, particularly those which it was proposed to introduce as parts of a deep-water, as distinct from a shallow-water, system. Nevertheless it was apparent that if means for laying and recovering the large repeater housings in deep water without damage to the cable were later developed the design would provide great flexibility in repeatered cable systems of the future. The authors therefore recommended that the British design should be used for the link between Newfoundland and Nova Scotia, where the water is of moderate depth and offers the opportunity for observation of a design based on proved shallow-water techniques. The design because of its later initiation, permitted of a lower cost per telephone circuit on this short link.

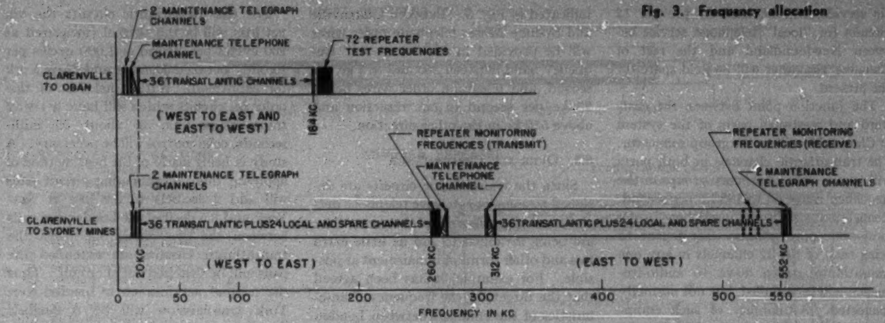
The Canadian Overseas Telecommunication Corporation was asked to join in the enterprise, because of its interest in overseas communication and because of the desire that the project should also improve United Kingdom-Canada communication. Administrative and technical discussions took place on both sides of the Atlantic, and agreements were arrived at and formalized in a contract signed on November 27, 1953. Ownership of the system will be vested in the American Telephone and Telegraph Company, the British Post Office, the Canadian Overseas Telecommunication Corporation, and the Eastern Telephone and Telegraph Company. The latter is a subsidiary of the American Telephone and Telegraph Company and will be concerned with the facilities within Canadian territorial limits.

6. General Description of the Complete System

6.1 Route

A map of the system is shown in Fig. 1, and the route in Newfoundland is shown in greater detail in Fig. 2. The route between Clarenville in Newfoundland and Oban in Scotland has been chosen so as to be well clear of existing telegraph cables, which lie to the south, and also to avoid known "holes" in the sea bed. Selection of a landing site in Newfound-

Fig. 3. Frequency allocation



land was determined by the following considerations:

1. The cables remain in deep water until they enter the northwest arm of Random Sound, where damage is not to be anticipated from grounding icebergs.

2. There is no crossing over other cables.

3. Clarenville provides reasonable amenities for the staff who will have to maintain equipment associated with the cables to Scotland and Nova Scotia.

Both the cables between Clarenville and Oban will be approximately 1,950 nautical miles long, and each will be equipped with 52 repeaters. The greatest depth at which a repeater will lie will be approximately 2,300 fathoms (2.6 statute miles), but most will be at depths between 1,200 and 2,000 fathoms.

Three alternative routes connecting Clarenville with Nova Scotia were considered; none was very attractive from an engineering standpoint. They were:

1. To leave Clarenville generally in an easterly direction and to lay the cable in a wide sweep round the Avalon Peninsula and in deep water to the south of the Newfoundland Banks.

2. To leave Clarenville generally in a southerly direction, cross the Avalon Isthmus (about 5 miles wide) into Placentia Bay, cross the Burin Peninsula (about 10 miles wide) into Fortune Bay, and thence in water of moderate depth to Sydney Mines.

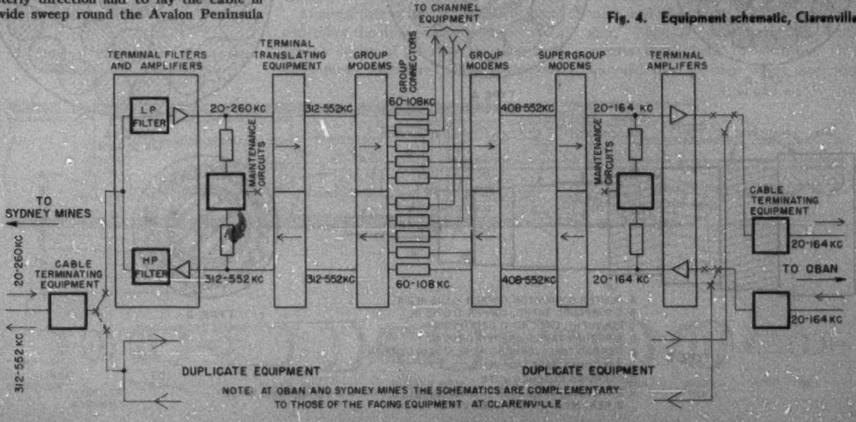
3. To lay the cable overland to Terrenceville at the head of Fortune Bay, thence picking up the same inshore route as in item 2.

The first of these would have involved a cable 600 nautical miles long crossing a number of existing telegraph cables; the second, while avoiding the main fishing grounds and other cables, would have necessitated landings at four intermediate points. Therefore, after an on-site investigation, the last alternative was chosen. This part of the system thus falls into two parts: to the east about 62 statute miles overland, and to the west about 274 nautical miles in coastal waters at depths up to about 250 fathoms. The single cable will be equipped with 16 repeaters, two of which will be on land.

6.2 CIRCUIT CAPACITY

Fig. 3 shows the location of the various frequency bands used for transmission and the location of various pilot tones. The cables between Clarenville and Oban will have transmission frequency bands extending from 20 to 164 kc per second and will provide capacity for 36 carrier telephone circuits, with channels at 4 kc per second spacing. The capacity of the cable between Clarenville and Sydney Mines will be somewhat greater, because it is planned to transmit the frequency band 20 to 260 kc per second from Sydney Mines to Clarenville, and the frequency band 312 to 552 kc per second in the reverse direction. The cable will therefore provide 60 carrier telephone circuits with channels spaced 4 kc per second apart. One of the two 12-channel groups in each direction, not employed for transatlan-

Fig. 4. Equipment schematic, Clarenville



TO CHANNEL EQUIPMENT
TERMINAL FILTERS AND AMPLIFIERS
TERMINAL TRANSLATING EQUIPMENT
GROUP MODEMS
GROUP CONNECTORS
GROUP MODEMS
SUPERGROUP MODEMS
TERMINAL AMPLIFERS
LP FILTER
HP FILTER
20-260 KC
312-552 KC
312-552 KC
312-552 KC
60-108 KC
60-108 KC
408-552 KC
408-552 KC
20-164 KC
20-164 KC
MAINTENANCE CIRCUITS
MAINTENANCE CIRCUITS
TO SYDNEY MINES
20-260 KC
312-552 KC
CABLE TERMINATING EQUIPMENT
CABLE TERMINATING EQUIPMENT
20-164 KC
20-164 KC
TO OBAN
DUPLICATE EQUIPMENT
DUPLICATE EQUIPMENT

NOTE: AT OBAN AND SYDNEY MINES THE SCHEMATICS ARE COMPLEMENTARY TO THOSE OF THE FACING EQUIPMENT AT CLARENVILLE tic service, will be used to provide 12 circuits for "local" telephone service between Newfoundland and the rest of Canada; the other will be held spare for the present.

The junction point between the eastward and westward parts of the system at Clarenville will be the group connector. The transatlantic channels in both parts will appear here as primary groups in the frequency range 60 to 108 kc per second. Connection will also be made here to the local and spare groups. Equipment to bring each of the 12 channels in any one transatlantic group down to audio-frequency will be installed, but not normally connected. Availability of such equipment is desirable in order to enable impairments affecting only one or two channels to be localized to one part of the system. Fig. 4 is a schematic of the arrangements at Clarenville. The arrangements will be similar at Oban and Sydney Mines.

The traffic operating terminals will be in London, New York, and Montreal. Twenty-nine telephone circuits will be put into service between London and New York and six between London and Montreal. Telegraph circuits between London and Montreal will be provided in the upper half of the 36th telephone circuit.

Telephone and telegraph circuits for maintenance purposes will be provided between Oban and Clarenville in the frequency band below 20 kc per second, as indicated in Fig. 3. Between Clarenville and Sydney Mines telephone order wires will be provided in the "crossover" frequency band between 260 and 312 kc per second, and telegraph order wires below 20 kc per second in one direction and above 552 kc in the other direction.

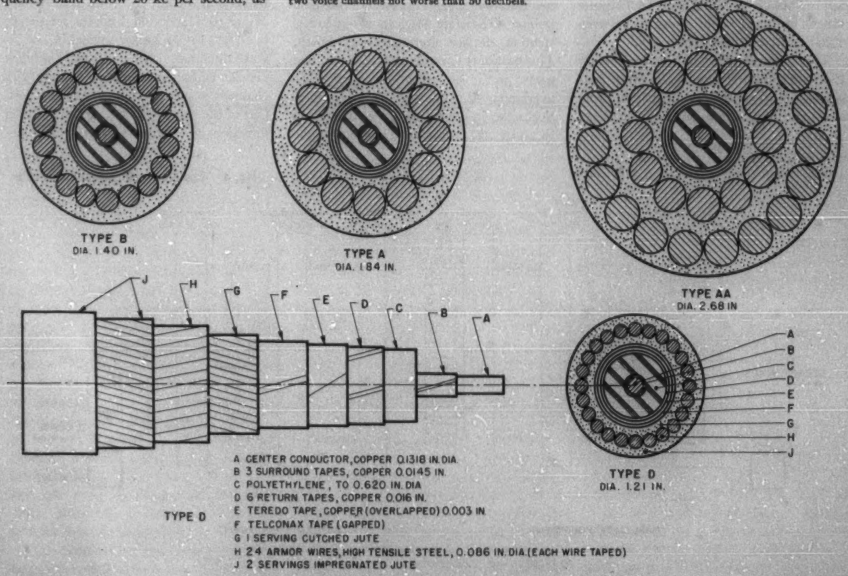
6.3 Over-all Performance

Since the transatlantic circuits are designed to connect extensive telephone networks on opposite sides of the ocean, they are being arranged to add as little extra loss and other forms of impairment as possible. For example, it has been agreed that the target for the frequency characteristics of the channels between London and New York, or London and Montreal, should be as good as that recommended by the Comité Consultatif International Téléphonique for a 2,500-kilometer international carrier circuit. Other objectives for the circuits compare favorably with those specified for long-distance circuits wholly on land.*

For the transatlantic circuits the via net loss will be 0.5 decibel (measured at 800 cycles per second or 1,000 cycles per second) between London and New York (or Montreal). With a net loss of this order on circuits which will have a 1-way transmission time of about 35 milliseconds, echo control will be necessary. A study is being made of the best method of applying this. Terminating circuit pads will add 2 decibels to the loss at New York and 3.5 decibels at London for calls between the two cities, but when the transatlantic circuits are extended the pads will be switched out of circuit. Thus the normal net loss on a London-New York transmission will be 6 decibels (3.5 + 0.5 + 2), or slightly more than the objective for Regional Center to Regional Center connections in the United States or Canada. It is an agreed objective to restrict variations from the specified net loss to a standard deviation of 1.5 decibels. This will put a severe requirement on the individual links which may be difficult to meet. In all respects the requirements for the Sydney Mines-Clarenville-Oban cable sections have to be somewhat more stringent than for the over-all circuits.

Vol.
Pt. 1
1955

Fig. 5. Types of cable


TYPE B
DIA. 1.40 IN.
TYPE A
DIA. 1.84 IN.
TYPE AA
DIA. 2.68 IN
J
H
G
F
E
D
C
B
A
TYPE D
DIA. 1.21 IN.
A CENTER CONDUCTOR, COPPER 0.1318 IN. DIA
B 3 SURROUND TAPES, COPPER 0.0145 IN.
C POLYETHYLENE, TO 0.620 IN. DIA
D 6 RETURN TAPES, COPPER 0.016 IN.
E TEREDO TAPE, COPPER (OVERLAPPED) 0.003 IN
F TELCONAX TAPE (GAPPED)
G 1 SERVING CUTCHED JUTE
H 24 ARMOR WIRES, HIGH TENSILE STEEL, 0.086 IN. DIA. (EACH WIRE TAPED)
J 2 SERVINGS IMPREGNATED JUTE
TYPE D

---

* The following have been agreed for the over-all circuits London-New York or London-Montreal:

1. Frequency stability better than ±2 cycles per second.

2. Noise at zero transmission level smaller than 46 decibels below 1 milliwatt, read on the Comité Consultatif International Téléphonique (1951) psophometer; or 38 decibels adjusted, read on the Bell system 2B noise measuring set with F1A weighting.

3. Minimum equal level crosstalk between any two voice channels not worse than 50 decibels.

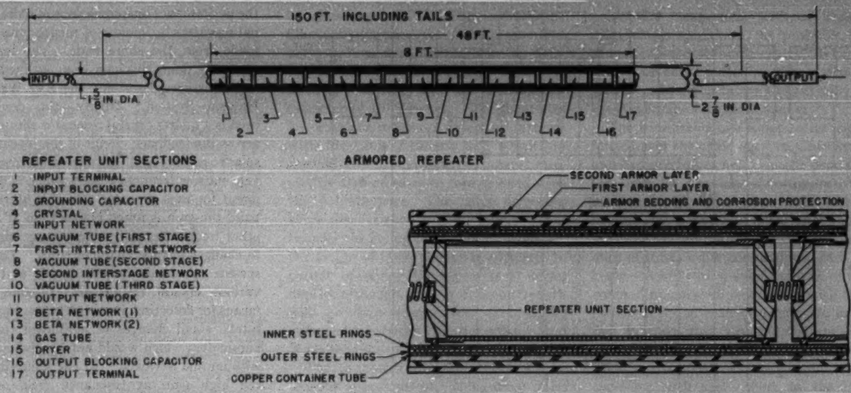

150 FT. INCLUDING TAILS
48 FT.
8 FT.
INPUT
OUTPUT
1 5/8 IN. DIA
2 7/8 IN. DIA
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
REPEATER UNIT SECTIONS
1 INPUT TERMINAL
2 INPUT BLOCKING CAPACITOR
3 GROUNDING CAPACITOR
4 CRYSTAL
5 INPUT NETWORK
6 VACUUM TUBE (FIRST STAGE)
7 FIRST INTERSTAGE NETWORK
8 VACUUM TUBE (SECOND STAGE)
9 SECOND INTERSTAGE NETWORK
10 VACUUM TUBE (THIRD STAGE)
11 OUTPUT NETWORK
12 BETA NETWORK (1)
13 BETA NETWORK (2)
14 GAS TUBE
15 DRYER
16 OUTPUT BLOCKING CAPACITOR
17 OUTPUT TERMINAL
ARMORED REPEATER
SECOND ARMOR LAYER
FIRST ARMOR LAYER
ARMOR BEDDING AND CORROSION PROTECTION
REPEATER UNIT SECTION
INNER STEEL RINGS
OUTER STEEL RINGS
COPPER CONTAINER TUBE

Fig. 6. Repeater structure, Clarenville-Oban section

A pilot tone at a frequency slightly displaced from 84 kc per second will be injected into each group of 12 channels at a point where the channels appear in the 60- to 108-kc-per-second range. Measurements of the received level of the tone will give an indication of the over-all net loss; later the received tone may be used to control automatically the gain of terminal amplifiers on each group. In addition, 92-kc-per-second pilot tones, similarly injected, will be transmitted over each cable section to give an indication of the net loss of the section. These tones will be blocked from passing from one section to another.

Signaling over the cable circuits will be designed to allow co-ordination between the different national long-distance signaling systems at the ends.

7. Clarenville-Oban Section

As already indicated, two cables will be used between Newfoundland and Scotland, one for each direction of transmission. These and the associated repeaters will be of the type now in operation between Key West and Havana, and have been described in some detail in the technical literature.[4]

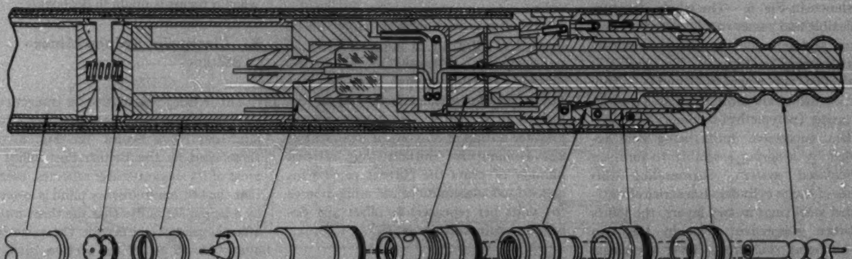
7.1 CABLE

The evolution of the cable design has been described in sections 2 and 3.4. The cable to be used on the transatlantic route will have a somewhat greater thickness of dielectric than that included in the make-up of the Key West-Havana cables, and will be 0.62 inch in diameter. (Cable dimensions refer to the core diameter before the addition of the outer conductor.)

Manufacture of the cable will demand control of the dimensional tolerances to a greater degree of precision than has been attempted heretofore. This is because the large number of repeaters in circuit make it of great importance to maintain a fine balance between repeater gain and cable loss. Predetermination of the effects of temperature and pressure on cable attenuation will enable this fine balance to be obtained when the cable is at sea-bottom temperature and pressure. Impedance irregularities must also be avoided, since they affect the gain-frequency characteristics of the repeaters.

As is customary in cable practice, special cable armoring will be used near the shore ends, and the various types of armoring are shown in Fig. 5. Type-*A* cable has an armoring consisting of 12 galvanized mild-steel wires, each of 0.3-inch diameter, and is laid in water not deeper than about 300 fathoms; type *AA*, used for the landing, is similar but has an extra layer of armoring. Type *B* is laid in moderate depths and has 18 galvanized mild-steel wires, each of 0.165-inch diameter. Type *D* is laid in all depths greater than about 700 fathoms. It is armored with 24 high-tensile steel wires

Fig. 7. Internal construction of repeater, Clarenville-Oban section


HIGH VOLTAGE CAPACITOR
COUPLER ASSEMBLY
END TERMINAL
GLASS SEAL ASSEMBLY
RUBBER SEAL ASSEMBLY
PROTECTIVE TUBE ASSEMBLY
COVER
NOSING PARTS
CORE TUBE

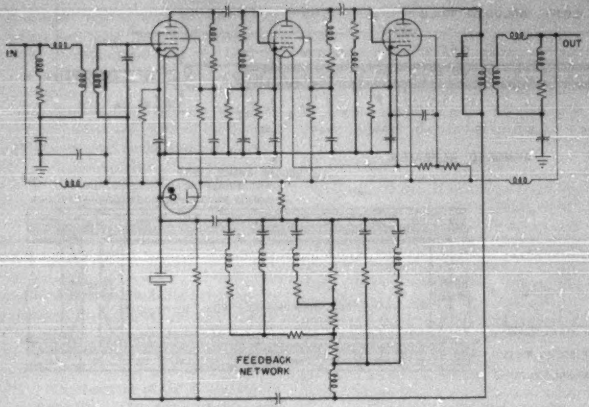


Fig. 8. Repeater schematic, Clarenville-Oban section

of 0.086-inch diameter, each wire being taped; these enable it to be laid and lifted in the deepest water.

7.2 Repeaters

The desirability of a flexible type of housing was referred to in section 3.1. The structure which was finally evolved is a flexible bulge in the cable 8 feet long and 2.9 inches in diameter, tapering, for a distance of about 20 feet at each end, down to the cable diameter. The armor of the cable itself is continued over the repeater housing, but with armor wires added to give complete coverage. To prevent the twisting of the repeater structure in the laying operation, there is a second layer of wires with opposite lay. The repeater itself is enclosed within an 8-foot-long series of cylindrical sections inside the housing.

The structure of the repeater container is shown in Fig. 6. The whole structure is flexible and passes readily over the cable drum and bow sheave without requiring that the ship be stopped. To attain this flexibility the repeater elements within the container are mounted inside a series of lucite (polymethyl-methacrylate) cylinders, successive units being held together by a spring assembly to form an articulated system. Surrounding this series of lucite cylinders is a series of butt-ended steel rings in two layers, the joints between successive rings in the two layers being staggered. Over the rings, and supported by them against collapse at sea-bottom pressure, is an envelope in the form of a long copper tube over which are the protective coatings and armoring wires

The repeater enclosure is terminated at each end with a system of seals, comprising first a glass-metal seal adjacent to the repeater elements, next a plastic seal molded to the cable insulation and, finally, at the extreme ends, a 7-foot-long seal formed within a copper tube which is an extension of the repeater housing. Sea water, in order to penetrate the housing, would therefore have to thread a long, multibarriered path. All of the seals are adapted for the sea-bottom pressures that they may have to withstand. A partially sectioned portion of the seal is shown in Fig. 7, together with further details of the repeater structure.

Fig. 8 is an electrical schematic of the amplifier. The lucite cylinders into which the electric elements will be packaged are indicated in Fig. 6. Excluding the tubes and high-voltage capacitors, 54 components will be accommodated within a volume no greater than that of a 3-inch cube.

The 52 repeaters in each cable will be located at intervals of approximately 37 nautical miles. They will be fed with d-c power originating at each shore end, the power feeding arrangements being described in section 9. Each repeater will produce a drop of 55 volts at a current of 225 milliamperes initially, but after a number of years the current can be increased to a maximum of 245 milliamperes (64 volts per repeater) to offset any deterioration in tube emission.

7.3 Testing Arrangements

Frequencies in the range 167 to 174 kc per second will be used to keep track of the performance and conditions of individual repeaters by means of measurements made from the shore ends. The techniques for doing this have been outlined in earlier articles describing the Key West-Havana System.[4] Briefly, the device used is a quartz crystal resonator shunted across the feedback circuit of the repeater so as to remove feedback effectively over a very narrow frequency band which is different for each repeater. In this narrow band the gain is much higher than at all other frequencies and is very susceptible to changes in tube characteristics. Thus system gain measurements, made at the various crystal frequencies, provide a means for detecting, in each repeater separately, small changes in significant tube parameters which would not be disclosed with normal feedback. At the same time the high gain at resonance creates a series of thermal noise peaks at the crystal frequencies. These can be observed by a narrow-band detector at the receiving end of the cable and will provide a means for locating repeaters which have failed. To accomplish this it is only necessary to note the missing noise peaks, since these are associated with the repeater which has failed and those preceding it. In order to make this test possible, even when the failure is caused by an open-circuit heater, a gas diode is provided in each repeater to by-pass the power-supply current.

7.4 Repair Repeaters

If the cable is damaged during laying or at a later time, it will have to be raised and repaired. After repair the cable length in the repaired section will be greater than before, and this will upset the balance between cable loss and repeater gain. In the deepest water the extra cable length may amount to several miles, and, to compensate for the additional loss, low-gain "repair repeaters" are being provided. One of these will be inserted when a repair is made in deep water.

8. Clarenville-Sydney Mines Section

The frequency bands to be transmitted over the single cable between Newfoundland and Nova Scotia (see Fig. 3) are those used by the British Post Office in most of its shallow-water schemes, except that the bottom frequency band is lowered by 4 kc per second so that the three transatlantic groups can occupy the same frequency band as on the section between Newfoundland and Scotland; this will simplify the frequency-translating equipment.

8.1 SUBMARINE CABLE

The structure and dimensions of the conductors and core of the submarine cable will be identical with those of the cable used between Clarenville and Oban; the cable will have *A*- or *B*-type armoring (Fig. 5). It will, however, be tested to a specification appropriate to the higher frequencies to be transmitted, and taking into consideration the fact that some slight easements are possible because the impedance of the repeaters in this section will approximately match that of the cable.

8.2 LAND CABLE

The land cable will be basically the same as the submarine cable but, being more subject to electrical interference, it will be shielded with layers of soft iron tape applied over the outer conductor. The tape in turn will be protected from corrosion by a polyethylene sheath, over which will be applied jute bedding and armor wires. The cable will be laid in a trench at a depth of approximately 30 inches, where the temperature is expected to have an annual variation of about 35 degrees Fahrenheit and a maximum daily variation of about 2 degrees Fahrenheit. Guard wires for protection of the system against lightning will be laid in the trench above the cable. The two repeaters in the land section will be identical with those in the sea and, according to the terrain determined in the final survey, will be laid in ponds, buried in the ground, or located in small manholes.

8.3 REPEATERS

The mechanical design of the repeaters which will be used in this section is shown in Fig. 9. The electric equipment is mounted in an inner sealed cylinder 4 feet 2 inches long and $7^{3}/_{4}$ inches in diameter, filled with dry nitrogen. This is mounted between brazed-in bulkheads in the pressure-resisting outer housing, which is about 9 feet long and has a maximum diameter of $10^{1}/_{2}$ inches. The glands are polyethylene moldings which are integral with the cable dielectric and which are shrunk on to castellated bosses, previously treated so as to form a bond with the polyethylene. The cable armoring wires are fastened under clamps which transfer the laying tension to the housing; the design of these clamps is adequate to permit of laying the repeaters at ocean depths, although, of course, such strength will be unnecessary for cable-laying operations between Newfoundland and Nova Scotia.

The 16 repeaters in this section will be located approximately 20 nautical miles apart. Electrically, they will be arranged to give 2-way amplification, as shown in Fig. 10. A feature of the design is the inclusion of balanced bridges at each end of the repeater. These not only supplement the directional discrimination in the filters by a minimum of 20 decibels but also contribute substantially to the equalization. The bridges also have the advantage that the intermodulation requirements in the filters are reduced so that small, high-*Q* ferromagnetic-cored coils can be used. Apart from the equalization provided by the directional bridges, the repeater derives most of the necessary gain/frequency characteristic from the 2-terminal feedback network.

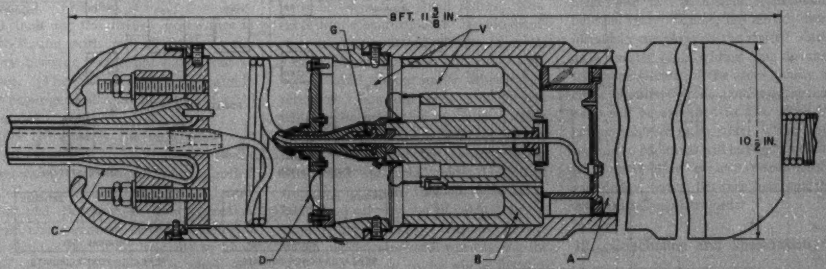
There are two 3-stage amplifier units in parallel between common input and output transformers with a single feedback path as shown in Fig. 11. The feedback network will maintain the repeater gain if one amplifier unit becomes inoperative, but the overload point is reduced by 5 decibels under this condition; a faulty component in one amplifier unit cannot affect the other unless it introduces circuit noise. In comparison with a system using single amplifiers, the probable life of each repeater is increased and the probability of early faults on the system is greatly reduced as shown theoretically in the Appendix of an earlier paper.[8]

The wide frequency range to be transmitted over the cable between Clarenville and Sydney Mines will be made practicable only by the use of comparatively modern high-performance tubes with a mutual conductance of 6,000 micromhos. It is well known that the mutual conductance of tubes using conventional cathode cores—nickel with a small content of silicon and magnesium—decreases with time owing to the growth of a resistive interface between this core and the emissive material of the cathode; this interface applies negative feedback to an extent which increases with age. Work in the Post Office Research Laboratories has shown that the formation of an interface can be prevented by the use of pure platinum cathode cores. Cores of this kind, known as "passive" cores, will be used; these increase the difficulties of processing, but result in tubes which should have stable characteristics over very long periods of time.[9]

The heater of each tube requires about 300 milliamperes at 5.5 volts, and the six heaters in each repeater are series-connected; the cathode-anode voltage (90 volts) is derived from a separate resistor carrying the line current. Each group of three heaters is bridged by a device which will short-circuit the group if a heater fails, and thus maintain the line current.

The space afforded by the rigid housing not only permits the complicated circuitry of the British repeaters but also makes it possible to design some of the components on more generous lines than can be fitted into the American flexible housings. For example, the power capacitors use a nonpolar dielectric of mineral oil instead of polar castor oil and are only

Fig. 9. Repeater housing, Clarenville-Sydney Mines section



A INNER UNIT ENCLOSING ELECTRICAL APPARATUS
B BULKHEAD (BRAZED IN)
C ARMOR CLAMP
D FLEXIBLE DIAPHRAM ENCLOSING VISTAC (V)
G CASTELLATED GLAND

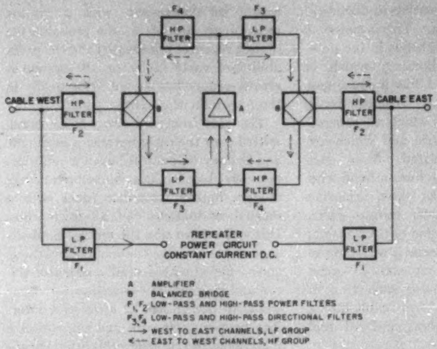

F4
F3
HP FILTER
LP FILTER
CABLE WEST
HP FILTER
F2
B
A
B
HP FILTER
CABLE EAST
F2
LP FILTER
F3
HP FILTER
F4
LP FILTER
F1
REPEATER
POWER CIRCUIT
CONSTANT CURRENT D.C.
LP FILTER
F1
A AMPLIFIER
B BALANCED BRIDGE
F1, F2 LOW-PASS AND HIGH-PASS POWER FILTERS
F3, F4 LOW-PASS AND HIGH-PASS DIRECTIONAL FILTERS
→ WEST TO EAST CHANNELS, LF GROUP
← EAST TO WEST CHANNELS, HF GROUP

Fig. 10. Schematic of 2-way repeater

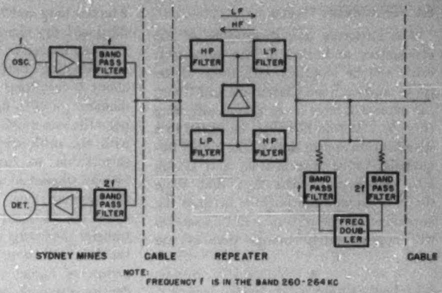

LF
HF
f
OSC.
f
BAND PASS FILTER
HP FILTER
LP FILTER
LP FILTER
HP FILTER
2f
DET.
BAND PASS FILTER
f
BAND PASS FILTER
2f
BAND PASS FILTER
FREQ. DOUBLER
SYDNEY MINES
CABLE
REPEATER
CABLE
NOTE:
FREQUENCY f IS IN THE BAND 260-264 KC

Fig. 12 (above). Repeater fault location

subject to about one-third the electric stress. Silvered mica capacitors will be used only where the polarization is less than 10 volts; at all other points the capacitors will be of the sealed oil-impregnated paper type. All components will be manufactured and tested to standards similar to those described in section 3.3.

8.4 Testing Arrangements

It is an advantage of a 2-way cable that, provided some form of frequency translation is possible in each repeater, signals transmitted from one terminal can be returned from the repeaters to the same terminal. In this instance the frequency translation in the repeater is effected, as shown in Fig. 12, by a frequency-doubler associated with filters accepting frequencies in the range 260 to 264 kc (i.e., from Sydney Mines) which are individual to the repeater. Signals at twice these frequencies are returned to the sending terminal and enable an accurate measurement to be made of the transmission level at the output of each repeater at any time.

Pulse-testing equipment will also be provided at each terminal to permit measurements of the overload point of each repeater. The repeaters will be identified on a time base and the frequency translation effected by the overload distortion of the amplifier itself. This test can be carried out in either direction: by direct harmonic production in one direction and by intermodulation of a 2-tone signal in the other direction. It is aimed primarily at determining whether both amplifiers are functioning in each repeater.

9. Power Feed to Repeaters

The arrangements for feeding power to the repeaters on the two sections are generally similar and involve the use of constant-current generators at both terminal stations, series-aiding and arranged so that one end controls the value of the line current; see Fig. 13.

The repeaters on the Clarenville-Oban section will require a total driving voltage of about 3,900 volts initially; this will be increased later to about 4,450 volts, or up to about 4,700 volts if it is necessary to introduce additional repeaters after repairs. The maximum voltage on the cable will be half of this. For the repeaters on the Clarenville-Sydney Mines section the necessary total driving voltage will be about 2,300 volts and, if necessary, it will be possible to feed this from one end only, since the power capacitors are adequately rated.

The constant-current generators used on the two sections will differ in several respects. Those on the Clarenville-Oban section will derive the constant-current feature by feeding via high-impedance pentode tubes with a mechanical regulator to absorb all but short-term variations. To assist continuity of the line current, the load will be shared by two generators at each end, each generator being capable of taking the full load. Transductor-controlled power units will be used for the Clarenville-Sydney Mines section.

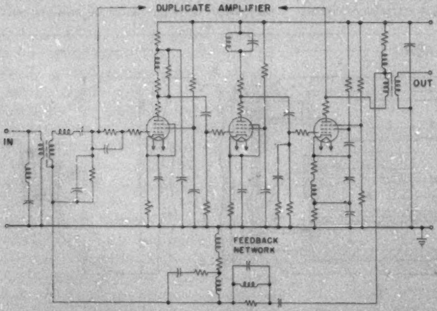

DUPLICATE AMPLIFIER
IN
OUT
FEEDBACK NETWORK

Fig. 11. Amplifier schematic, Clarenville-Sydney Mines section

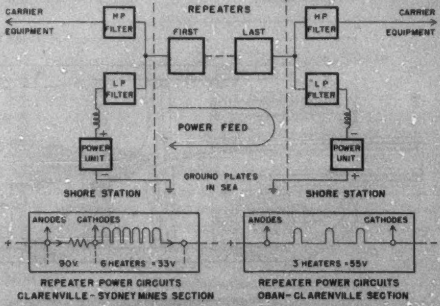

CARRIER
EQUIPMENT
HP FILTER
REPEATERS
FIRST
LAST
HP FILTER
CARRIER
EQUIPMENT
LP FILTER
LP FILTER
POWER FEED
+
POWER UNIT
−
−
POWER UNIT
+
GROUND PLATES IN SEA
SHORE STATION
SHORE STATION
ANODES
CATHODES
90V
6 HEATERS = 33V
ANODES
CATHODES
3 HEATERS = 55V
REPEATER POWER CIRCUITS
CLARENVILLE - SYDNEY MINES SECTION
REPEATER POWER CIRCUITS
OBAN - CLARENVILLE SECTION

Fig. 13. Basic power feeding arrangements

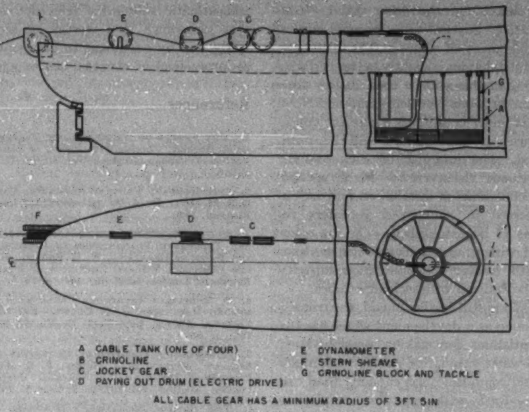

F
E
D
C
G
A
B
A CABLE TANK (ONE OF FOUR)
B CRINOLINE
C JOCKEY GEAR
D PAYING OUT DRUM (ELECTRIC DRIVE)
E DYNAMOMETER
F STERN SHEAVE
G CRINOLINE BLOCK AND TACKLE
ALL CABLE GEAR HAS A MINIMUM RADIUS OF 3FT. 5IN

Fig. 14. Cable-laying machinery on HMTS Monarch

Every effort will be made to ensure continuity of the d-c power supply to the cable so as to avoid voltage surges at the repeaters and, in order to assist this the a-c supplies to each power feed unit will be derived via a-c–d-c–a-c machines. These machines will be driven normally from commercial power supply mains, but in the event of failure diesel-driven alternators will be started up automatically. There will be two stages of stand-by:

1. The a-c–d-c–a-c machines will continue to run, driven from the 130-volt battery which energizes the d-c machine.
2. If the commercial power failure is of long duration the stand-by diesel plant will take over.

10. Cable-Laying Arrangements

All the cable will be laid by HMTS *Monarch*. This ship, of 8,056 gross tons, was built for the British Post Office in 1945,[10] and with full oil bunkers can carry between 5,000 and 6,000 tons of cable. (The tons referred to here, and in the paper generally, are long tons of 2,240 pounds.)

The *Monarch* is the only cable ship afloat capable of laying the whole of the deep-water part (about 1,500 nautical miles) of each cable between Newfoundland and Scotland in one operation. From earlier sections of the paper it will be obvious that it is desirable to avoid any sequence of operations which will make it necessary to pick up a cable end and to make a splice in deep water. Such operations will have to be carried out if repairs are necessary after the cable has been laid, but inevitably introduce some risk of the cable kinking; the initial laying will be planned to avoid them. All repeaters will have been jointed into the cable before it is loaded into the ship.

A period of 12 days, during which continuously good weather conditions may be anticipated, is necessary in order to carry out the main operation without undue hazard to the repeatered cable. In the North Atlantic such conditions are unlikely except during the period from mid-May to mid-September, and it will not be practicable to lay both the cables on the main crossing in one summer. It is planned to lay one of these cables during the summer of 1955 and to complete all cable-laying operations in 1956.

The essential parts of the after cable-paying-out machinery are shown in Fig. 14. Although the flexible, built-in repeaters are designed to be handled as part of the cable, it is of course desirable not to bend them to a small radius. New cable drums, approximately 7 feet in diameter, are being fitted on the *Monarch*. This is the largest size that can be accommodated without major structural alterations. Bow and stern cable sheaves will also be made this diameter. Roller guides with a minimum radius of 3 feet 6 inches will replace the existing spider wheels and "spectacle" guides throughout the ship to ensure freedom from sharp bends and an easy run of cable from the tanks during laying. Various devices which proved successful when laying the Key West-Havana cables will be fitted to ensure that the repeaters conform to the curvature of the cable drums and to ease their passage round the drums and at other points.

Three new dynamometers will also be fitted, two forward and one associated with the paying-out aft gear. Seven-foot-diameter sheaves will be used, associated with electrically sensitive pressure cells. The latter will be arranged to operate indicating and recording instruments on deck and in the test room, and will enable the tension on the cable to be closely controlled as it is paid out.

The laying of the cable between Newfoundland and Nova Scotia will present no special problems, but an electrically driven hoist has been fitted over the ship's bows to facilitate handling the rigid repeaters.

11. Connection of Cable System to American, Canadian, and British (and European) Internal Telephone Networks

A description of the cable and radio-relay systems which will be used to extend the transatlantic circuits from Oban to London and from Sydney Mines to New York and Montreal would be outside the scope of this paper and, in any case, systems of the same kind have already been used extensively in the countries concerned. It may be of interest, however, to indicate how the connections will be made. From Oban the transatlantic circuits will be taken to London by carrier cable; there will be two alternative routes, one via Glasgow and the other via Inverness and Aberdeen. In London it will be possible to connect any transatlantic circuit to any one of the existing submarine cable circuits to the continent of Europe, thus providing through cable communication between North America and continental Europe. On the American side a microwave radio-relay system will connect Sydney Mines with Portland, Maine. This type of system, which makes use of line-of-sight transmission between relay stations, has already been installed extensively in the United States as a means of providing long-distance telephone circuits of the same quality as cable circuits. From Portland the circuits will be extended to New York by carrier cable or radio relay. Cable connection to Montreal will be effected either at Sydney Mines or at a point on the radio-relay system between here and the American-Canadian border.

12. Conclusion and Speculation

Barring unforeseen and unexpected delays, 1956 should see the completion of the first transatlantic telephone connection via submarine cable. With the completion of this project, telephone service between the two continents should enter into a new era marked by improved quality of service and reliability and a capacity for growth no longer restricted by the limited capacity of the radio spectrum. The advantages of radio will remain, however: flexibility, the speed with which communications can be established and shifted from route to route, the feature of direct access from one country to another, relatively low cost, etc. These are not inconsiderable advantages and they augur for the continued importance of short-wave radio telephone services.

On the other hand, the new transatlantic cable will by no means represent all that could be done if full advantage were taken of contemporary techniques. The submarine cable art is a conservative one. As Dr. Buckley observed, the laying of a submarine cable, like any other activity on the high seas, must be planned with its unavoidable hazards in mind. Experience over 100 years has shown that when handling deep-water submarine cables on the high seas it is better to be safe than sorry.

But the consequences of this doctrine of caution are equally plain. The submerged repeaters in the 1956 deep-water cable section incorporate tubes whose proven design dates back to 1941 and with a mutual conductance of 1,000 microhms. On the other hand, "conservative" postwar tubes provide mutual conductances of about 5,000 microhms, and tubes having a slightly better performance than this will be used in the cable system between Newfoundland and Nova Scotia under conditions where the consequences of tube failure will be minimized. The integrity of such postwar tubes under the severe and highly specialized conditions of installation and use encountered in submerged repeaters remains to be established. On the other hand, without question cable systems having still greater communication capacity can be realized with such new tubes.

Looking still further ahead, the transistor looms as a development which has the potentialities for making possible long, deep-water submarine cables with much greater communication capacity than can be realized with repeaters employing thermionic tubes. The voltage required by the tubes sets an upper limit on the number of them that can be operated in tandem before the power supply results in voltages on the cable that are beyond the limits of safety. The much lower power drain of the transistor should overcome this obstacle and permit more repeaters to be used and wider frequency bands to be accommodated. Ruggedness, long life, and small size are added attractive features of the transistor which are of not inconsiderable importance to the future development of submarine cables. A transatlantic submarine television cable is a long-range goal worthy of serious study and by no means to be dismissed as impractical of eventual attainment.

References

1. The Future of Trans-Oceanic Telephony, **O. E. Buckley.** *Journal*, Institution of Electrical Engineers, London, England, vol. 89, pt. I, 1942, p. 545.

2. One Hundred Years of Submarine Cables, **G. R. M. Garratt.** *H. M. Stationery Office*, London, England, 1950.

3. Fifty Years' Development in Telephone and Telegraph Transmission in Relation to the Work of Heaviside, **W. G. Radley.** Heaviside Centenary Volume, Institution of Electrical Engineers, London, England, 1950, p. 76.

4. A Submarine Telephone Cable with Submerged Repeaters, **J. J. Gilbert.** *Bell System Technical Journal*, New York, N. Y., vol. XXX, 1951, p. 65.

5. Submerged Telephone Repeaters for Shallow Waters, **R. J. Halsey, F. C. Wright.** *Proceedings*, Institution of Electrical Engineers London, England, vol. 101, pt. I, 1954, p. 167.

6. The British Post Office Standard Submerged Repeater System for Shallow-Water Cables with Special Mention of the England-Netherlands System, **D. C. Walker, J. F. P. Thomas.** *Ibid.*, p. 190.

7. The Netherlands-Denmark Submerged Repeater System, **A. H. Roche, F. O. Roe.** *Ibid.*, p. 180.

8. Modern Submarine Cable Telephony and the Use of Submerged Repeaters, **R. J. Halsey.** *Journal*, Institution of Electrical Engineers, London, England, vol. 91, pt. III, 1944, p. 218.

9. The Life of Oxide Cathodes in Modern Receiving Valves, **G. H. Metson, S. Wagener, M. F. Holmes, M. R. Child.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 99, pt. III, 1952, p. 69.

10. HMTS Monarch, **A. J. Gill.** *Post Office Electrical Engineer's Journal*, London, England, vol. 39, 1947, p. 129.

◆

Discussion

Oliver E. Buckley (retired) (Maplewood, N. J.): The authors have kindly invited me to open the discussion of the paper, and I welcome the opportunity both because of the great importance of the undertaking described and because of a deep interest I have long had in the subject.

The system described represents, in my view, the most radical and important advance yet undertaken in the old and conservative art of transoceanic cables. One way to measure this advance is the top frequency which can be transmitted. It took more than half a century to go from 1 cycle per second or so of the first transatlantic cable to the 100 cycles per second or so of the continuously loaded telegraph cables. Now it is planned to go in one great step to a top frequency of over 100,000 cycles per second and to transmit a broad band of frequencies to provide 35 or more telephone circuits.

One reason for the traditional conservatism of the cable art is the great cost in time and money of making repairs once a cable is laid. This has put a high premium on ruggedness and simplicity. In this new project, simplicity will be sacrificed to no small degree by incorporating in the cable structure a complicated assembly of electronic gear precisely designed to compensate for the characteristics of the cable over a wide range of frequencies, and there will be not just one such repeater but over 100 of them in the two cables for the Newfoundland-Scotland link. This calls for unprecedented care and precaution to ensure success. It has called for extraordinary devices and for imagining all possible causes of trouble that may show up and inventing means to deal with them.

One method of making such a repeatered cable was proposed in my 1942 Kelvin Lecture (reference 1 of the paper). Repeaters of the type there described, with important improvements subsequently added, will be used in the long and deepsea section of the system. The link to connect Newfoundland and Nova Scotia will employ 2-way repeaters developed by the British Post Office. Experience to date with installations of both types of repeater has demonstrated their practicability.

So radical a venture can only be justified by great value to be gained and that value is indeed very great in the present instance. Though transatlantic telephone service, as it is now provided by short-wave radio telephone systems, has become commonplace, it is still short of the perfection of long-distance overland telephone transmission of the coaxial cable and radio-relay systems. Further, overseas short-wave radio is, on occasion, subject to impairment or complete interruption by magnetic storms. Even when radio circuits are working well, there are not enough channels available in the short-wave spectrum to satisfy the ever-growing demand for public service.

The provision of good wire service across the Atlantic will, I feel sure, lead to greater public demand. This will in turn lead to more cables and I expect cables will become the backbone of transatlantic telephone communication.

Confident as I am of the practicability of repeatered submarine cables, and the need for them, I am equally sure that the present designs are not the ultimate and that still better cables will follow. Just as this first edition of a broad-band transatlantic telephone cable system has been made possible by new materials and techniques originated in the laboratory, so will improvements, many of which are now under way, come into being.

Making the present project a joint undertaking brings to it the products of two similar but different lines of development. This is significant of the spirit that has long characterized co-operation between Bell System engineers and the engineers of the British Post Office. Out of this co-operation will come benefits that will contribute to future progress.

The authors of this paper and all who

participated in the project are to be congratulated, but I think they will prefer to receive their congratulations after the job is done and they can talk about it over a circuit that will bring the speakers on both ends as close together acoustically as if they were in the same room.

Col. Sir A. Stanley Angwin (Commonwealth Telecommunications Board, London, England): Having been greatly interested in the first development of transatlantic telephony by radio and in the first proposals for a submarine telephonic cable, I am pleased to have the opportunity of commenting on the project outlined in this paper. It is most fitting that this example of Anglo-American co-operation in the field of telecommunications should be presented in a joint paper before the British Institution of Electrical Engineers and the AIEE.

This paper, with its picture of the growth of the North Atlantic wireless system and account of the first work on submarine repeaters, is of particular interes to the Institution as a corollary to the 33rd Kelvin Lecture given by Dr. Buckley (reference 1 of the paper), and I gave some indication of what was being done and contemplated in the United Kingdom in my Presidential Address of 1943.[1] (That address was, of course, made with many restrictions as to what could be publicized while the war was in progress.)

The tremendous impetus given to long-distance wireless from 1928 on stimulated the cable interests into consideration of the possibility of long-distance submarine cable telephony. An indication of the progress made is the extension of the bandwidth from that of the first transatlantic cable, 1.5 cycles per second, to 100 cycles per second, on the continuously loaded telegraph cables, and to 144 kc per second on the new telephone cable.

The conclusion has now been reached that the ideal should be the judicious combination of radio and cable for transoceanic conditions. In this particular case, not only will telephony by cable be added to radiotelephony between the United Kingdom, Canada, and the United States, giving a much more efficient service at all times, but the possibility is added of telephony from the United Kingdom to Australia and New Zealand over the cable and the trans-Canada line network and thence by radiotelephony on the Pacific route, thus improving the through telephony service to the antipodes to an extent that it will operate over those periods when it is impossible on the direct radio link. The authors are to be congratulated on presenting details of the design of the cable and estimates of its performance before the work is completed.

I think one of the outstanding factors that may be deduced from this paper is that the mechanical problems which have to be met in this undertaking are more critical than the electrical ones. Fundamentally, the solution of the repeater is similar in the United Kingdom and in the American design; the greater difference is in the housing, one articulated and flexible, the other bulkier and rigid.

There is a feeling of confidence in the design on the electrical side, a confidence inspired by exhaustive research and experiment, without which this great work could not have been undertaken. On the mechanical side, the conditions to be met are less determinate and the risks difficult to cover by a factor of safety; this is, perhaps, the most difficult feature and one to which much thought and trial have been given. Other problems are more intractable: the difficulties of cable laying in the Atlantic at depths down to 2,300 fathoms are touched on in the paper.

There have been many epics in the past of cable laying and cable repairs, from those of the laying of the first transatlantic telegraph cables, and as a more recent example of repairs, the attempt to restore the Porthcurno-Harbour Grace cable in 1944. In this attempt some 36 "drives" were made by the cable ship in a period of 40 days in conditions of fog and storm and other difficulties which prevented successful repairs; indeed restoration was not finally carried out until 1952.

I feel that when the first transatlantic telephone cables have been "well and truly laid" great credit will be due to the seamanship of those who overcome these hazards of nature. I hope this phase of the undertaking will be recorded and in due course brought before the Institution.

REFERENCE

1. INAUGURAL ADDRESS, Col. Sir A. Stanley Angwin. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 91, pt. I, 1944, p. 15.

H. J. Rosenberger (All American Cables & Radio, Inc., New York, N. Y.): In section 3.4 of the paper the statement is made: "The composite central copper conductor consists of a central solid wire surrounded by a single layer of spiralling and abutting tapes carefully formed to fit closely about the wire with a minimum of voids. Such a composite structure has better flexibility and protection against breakage than a solid conductor, and has better a-c and d-c resistance than a stranded-wire conductor of the same outside diameter."

It may be of interest to the authors to know that between 1943 and 1954 my company used approximately 500 nautical miles of solid core cable for cable repair work in depths up to 2,600 fathoms and in various lengths up to 24.5 nautical miles. Approximately 400 nautical miles of this had anhydrex rubber insulation and 100 nautical miles were insulated with a polyethylene compound. There has been no significant difference between the incidence of conductor breaks in this cable and in cable using stranded conductors. This repair cable used a 0.095-inch solid copper conductor insulated to a diameter of about 0.26 inch. It had no coaxial return.

Kenneth S. Wyatt (Phelps Dodge Copper Products Corporation, New York, N. Y.): The large amount of research in different fields, the long and intense effort, and the thoroughness of detail which have gone into the solution of the problem of transatlantic undersea telephone cables command respect and admiration. Once the cable is installed and operating, this paper will become a classic. It and its authors will long be remembered.

The authors point out that future cables will be still more efficient when contemporary techniques are more fully tried and advantage can be taken of them. The authors consider chiefly the repeaters and their components, but I think the efficiency of the cable can be greatly improved also. Better to use more efficient cables and fewer repeaters; in fact, on short runs of 200 nautical miles it is better to use an efficient cable and no repeaters at all. This is now feasible.

One way to design a more efficient cable is obviously to employ an air dielectric with a minimum of solid material spacing the inner from the outer conductor. The button or beaded cable is an example; its mechanical weakness in bending and crushing would render it completely impossible for undersea work; electrically it is undesirable for long lines because each button causes a reflection which is additive. However, by employing a helix made up of flexible polystyrene tapes, a wrap over the helix of polystyrene tapes, and an outer conductor and sheath of aluminum, a highly efficient, mechanically strong, economic cable has been attained. Tests show that such a cable is more than adequate mechanically for the laying operation, for the helix gives continuous support, and the styroflex insulation, being harder than the soft copper or the aluminum conductor, does not flow under the pressures involved. Greater precision of centering is obtained by applying the insulation in tape form rather than by extruding. The pitch of the helix is absolute. The electrical efficiency of such a cable derives from:

1. The lower attenuation in the air dielectric and the lower attenuation in the copper conductor (because the lower capacitance permits the use of a larger copper conductor for the same optimum impedance).
2. The effectively 1-piece construction reduces the noise level to the thermal noise level.
3. The extreme precision of the cable makes for negligible impedance variations and phase distortion.

With regard to the aluminum sheath, it is protected by an extruded jacket of polyvinyl chloride or polyethylene over which the standard armoring and hessian servings are applied.

To prevent flooding of the cable in case of damage, nonreflecting pressure dams are built into the cable as manufactured at intervals of 4 to 6 to the nautical mile. Allowing 115 decibels' total attenuation, made up of a signal input of +20 decibels referred to 1 milliwatt and a signal output of −95 decibels referred to 1 milliwatt, a styroflex cable of only 0.86-inch aluminum sheath inside diameter will carrying 60 voice channels of 4-kc spacing, starting at 12 kc and going up to 252 kc, a distance of 104 nautical miles without repeaters. The economic limit for such a system for civilian purposes would appear to be about 200 nautical miles.

There is a great need for an efficient undersea cable for carrier telephone without repeaters between islands. This is particularly true for defense purposes, where the unproved record and limited experience and possible high maintenance of repeaters are serious matters. I proposed such a solution of the communication problem between islands 100 miles or so apart, with shallow water in between, using land-type styroflex cable with the usual submarine corrosion

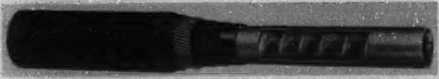


Fig. 15 (right). Shallow-sea styroflex cable. A highly efficient coaxial cable for carrier frequency voice transmission making repeaters unnecessary for distances up to 200 nautical miles. This cable is about 1.2 inches diameter inside the aluminum sheath and will handle 48 voice channels of 4-kc bandwidth for 200 nautical miles. Nonreflecting plugs are inserted every 1,000 feet to prevent flooding in case of damage

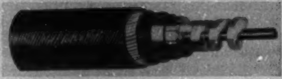


Fig. 16. Deep-sea styroflex cable. A special aluminum alloy armor makes the high-efficiency styroflex cable useful at ocean depths of 30,000 feet. With this cable the number of repeaters on transatlantic telephone circuits may be reduced to one-half or one-third, and possibilities of transatlantic television are thus opened up. Nonreflecting dams are inserted at intervals of 4 to 6 per nautical mile to prevent flooding in case of damage

protection and armoring, to the Canadian Government's Department of Defense Production in May 1952. The cable is shown in Fig. 15.

It is quite feasible, and possibly economical, to span the Atlantic with telephone cable without any submerged repeaters at all, and the authors are doubtless aware of this. In 1952 Halstead published his NARCOM plan[1] of connecting Europe and America with microwave and very high frequency by island-hopping over the North Atlantic, and he brought out the fact that the biggest interisland gap was only 290 nautical miles. I immediately saw that, although Halsted's plan was impracticable because of snow and ice conditions, the Nova Scotia - Newfoundland - Labrador - Greenland - Iceland - Faroe Islands - Shetland Islands-Scotland route was practicable with submarine carrier telephone cable if only the high-efficiency styroflex cable could be reinforced for the deep sea. In June 1952 Felten and Guilleaume of Cologne agreed to undertake a research program to develop such a cable, and within a year a construction was arrived at, after weeding out ideas from both sides, which fulfilled all requirements, including cyclic pressure tests up to 19,000 pounds per square inch. No commercial insulation other than styroflex could have been used. The deep-sea styroflex cable is shown in Fig. 16. A plan utilizing this cable for spanning the Atlantic via the North Atlantic Island Bridge without repeaters was submitted to the U.S. Air Force in June 1953. It carried with it a 5-year guarantee from one of the world's four large submarine cable makers.

The aforegoing discussion is not in any way theoretical. The authors have tested six repeaters on the Havana-Key West run for 4 years. The shallow-sea styroflex cable, it recently came to my attention, has been operating between Norway and Denmark for 14 years without incident. It handles 30 channels of 4-kc spacing (15 in each direction) over a distance of 80 miles. No repeaters are used. At places the water depth is over 2,000 feet. The cable has a lead instead of an aluminum sheath, and long-lay copper tapes under the lead serve as the return conductor.

It is my feeling that the most efficient methods are those with the utmost simplicity. If we can get rid of some or all of these complicated though ingenious submarine repeaters by the simple expedient of increasing the cable efficiency, the transoceanic submarine cable art will be better off. It is of course realized that there are economic limits to this proposal. These proposals in no way detract from the admiration and respect we hold for the repeater development described by the authors.

Reference

1. Many Uses for NARCOM Relay System, William Halstead. *Tele-Tech and Electronic Industries*, New York, N. Y., vol. 11, Oct. 1952, pp. 66–67.

J. J. Gilbert (Simplex Wire & Cable Company, Newington, N. H.): The authors of the paper and the eminent engineers who presented discussions have clearly indicated the difficulties and risks involved in the repeatered transatlantic telephone cable. Success will only be attained by adhering closely to the accumulated engineering experience and by use of the utmost care in the manufacture of repeaters and cable. Cable manufacture will be of unprecedented difficulty. Perhaps the most interesting requirement on cable performance is a consequence of the need for impressing potentials to ground of the order of 2,000 volts on the cable insulation in order to supply power to the repeaters. Until comparatively recent years it was considered unsafe to employ much more than 100 volts on long submarine cables. This restriction was based in part on sad experience in applying a high voltage to one of the early transatlantic cables, but there was also the realization that joints in insulation were not as strong electrically as the insulation itself. In the polyethylene cable, jointing has been studied intensively with the aim of obtaining a molecular bond rather than mere physical contact between the surfaces to be joined. The joint is then as strong and as lasting as the insulation itself. To gain still further assurance, the number of joints is reduced to a minimum by insulating the conductors in unusually long lengths.

Assurance of permanence and continuity of insulation and conductors presents many problems to the cable manufacturer; and there are others relating to the need, indicated in the paper, of achieving throughout manufacture predetermined values of electrical characteristics to an unusually high degree of precision and uniformity.

I would like to add to the tribute that has been paid to Dr. Buckley. All who have had contact with him in submarine cable development appreciate the inspiration of his leadership and the skill with which he placed this work on a firm foundation.

Mervin J. Kelly, Sir Gordon Radley, G. W. Gilman, and **R. J. Halsey:** We have been honored by the discussions from Sir Stanley Angwin and Dr. Oliver E. Buckley to whom the submarine cable art has owed so much for so many years. We express our thanks for their comments on the background of this project and for their confidence in its success. When the first transatlantic cable has been "well and truly laid," a complete discussion of the problems can certainly be expected.

Regarding Mr. Rosenberger's comments on the make-up of the center conductor, it is generally agreed that a solid conductor would be preferable for transmission reasons. Modern brazing techniques and methods of manufacture and inspection of the center conductor increase the likelihood of obtaining a mechanically dependable solid conductor. In the past, the use of soft soldered joints and the possibility of hidden inclusions in the copper have led to the choice of a composite structure and the present project has been based on designs which have proved their reliability by long usage.

Mr. Wyatt's comments can be divided into suggestions on the cable make-up and suggestions concerning the optimum cable route. With regard to his suggestion that nonrepeatered cables with partially air-spaced dielectric are best for distances less than 200 miles, the British Post Office has had 8 years' experience of such air-spaced polythene cable in the North Sea, providing, e.g., 84 circuits over a distance of 82 nautical miles. While these cables have been quite satisfactory, this type will not be used in the future because experience has proved that the circuits can safely be provided at much less cost by smaller diameter solid dielectric cables with repeaters.

On the transatlantic cable make-up, many different designs were studied. The main advantage of the design Mr. Wyatt describes is its lower dielectric constant which reduces attenuation for a given diameter. From Mr. Wyatt's figures, a 0.86-inch styroflex cable would have an attenuation of about 1.11 decibels per nautical mile at 252 kc per second, whereas a corresponding solid polyethylene dielectric cable of this size would have about 1.45 decibels per nautical mile at that frequency. On this basis, the increase in repeater spacing made possible by the proposed cable design would be about 30 per cent. If a much larger diameter styroflex cable were practicable for use in deep water, the ratio of the repeater spacing with such a cable to that with the present 0.62-inch solid polyethylene dielectric cable might be of the order of 2 or 3 to 1, as mentioned by Mr. Wyatt. In any case, the decision on the type of cable for this installation was made on the basis of proven integrity over a long period of usage and the solid dielectric cable was the only one with such a history in deep water.

On the matter of the optimum route, everyone interested in transatlantic cables has looked at a northern route—Shetlands, Faroes, Iceland, Greenland, Labrador (or Newfoundland)—with interest, and a number of detailed plans have been investigated, the first following the failure of the 1858 cable. The maximum distance between landings would be less than 1,000 nautical miles and the eastern part of the route would be easy. The western sections, however, would be difficult to lay and maintain in view of the ice conditions. Cable could usually be laid between Greenland and Newfoundland only in the months of September, October, and November when the days are short and the weather is poor, and year-round maintenance would only be possible if an ice-breaking cable ship were used. There is also considerable hazard from grounding icebergs. For a first transatlantic telephone cable, therefore, this route did not appear very safe.

We are also grateful for the comments by Mr. Gilbert who took a considerable part in the effort during the years before this project was launched.

A New Full-Wave Magnetic Amplifier Output Stage

P. W. BARNHART
ASSOCIATE MEMBER AIEE

ol. 74
. 1
955

THOSE concerned with the development of high-performance magnetic amplifier servo systems have long recognized the need for an output stage which would produce a full-wave current output into a servomotor and still retain the two inherent advantages, namely, the high speed of response and demodulation of the half-wave amplifier. In this paper a new concept in full-wave push-pull output stages is presented which fulfills this need. Test data are given to show its operation in an actual system.

As noted in an earlier paper on the half-wave bridge type magnetic amplifier,[1] the half-wave input to the motor contains a fundamental component which provides considerable torque. However, it is less than rated torque. The system designer must bear this in mind when he chooses a motor, because he will not be able to produce full-torque output from the motor even when using a shunt capacitor across the control phase windings. In general, the d-c component causes some loss of torque, about 15 per cent, but saturation effects vary widely from motor to motor. In many instrument servos this problem is not critical, but in other systems a better utilization of the servomotor is desirable.

While a conventional full-wave magnetic amplifier will provide the desired type of control current for the motor, it does not provide the two inherent advantages gained by using half-wave amplifier techniques, namely, speed of response and inherent demodulation. While much has been claimed about the speed of response of full-wave amplifiers with the relatively low gains used in servo systems, no one has produced a control system with such an amplifier that demonstrates the response of similar systems using half-wave amplifiers. The use of inherent demodulation characteristics to provide system compensation, as demonstrated by the U. S. Naval Ordnance Laboratory,[2] is not readily available when full-wave amplifiers are used. The use of compensation with full-wave amplifiers is usually concerned with improving amplifier response, not system response. Neither of these types of amplifier is the complete answer in the case where it is desirable to have a high-performance control system and utilize the servomotor to a maximum.

General Solution of the Problem

The ideal solution in this case is a half-wave amplifier in which the output can be converted to a full-wave voltage applied to the load. A circuit which will store information about the output of the half-wave stage into the load and produce the same output into the load on the next half-cycle will solve this problem. It must produce a phase-sensitive output with a minimum of d-c component, or a polarity-sensitive direct current.

The circuit which will most easily accomplish this task is simply another half-wave magnetic amplifier stage slaved to the output stage. A circuit to achieve this effect in single-ended amplifiers has been described by Lord.[3] Fig. 1 shows diagrammatically how such a "slave" stage operates. The input to the normal half-wave output stage is amplified and fed to the load. This amplified signal is also used as an input to the slave half-wave stage which, in turn, produces a like output to the load on the next half-cycle. The method of utilizing the output of the driver stage to control the slave stage and thus produce the desired full-wave output is very simple and can be accomplished in several ways.

Slave Stage Control in the Load Windings

The first method investigated was that of introducing control into the cores of the slave stage via its load windings. It had been noted early in the work on half-wave amplifiers that an alternating voltage across the load during the reset half-cycle of the output stage has a tendency to produce control on that stage if the rectifiers did not present a very high impedance. Thus, if the load on the amplifier were reactive, as with a servomotor, and the rectifiers did not have high back impedances (or were shunted for reset), the output stage would receive control from the reactive voltage across the motor and continue to produce output for several cycles after the signal was reduced to zero. Reference to Fig. 2 will show how this control is affected. This is the load circuit of a simple half-wave bridge output stage during the reset half-cycle when there was an output the preceding half-cycle. The voltage appearing across the load, e_0, and the line voltage e_L have the polarities shown. Instead of each load winding and its respective rectifier now having the voltage $e_L/2$ applied to it (as with a resistive load), *ab* and *ef* now have $(e_L+e_0)/2$ applied to them and *bc* and *de* have $(e_L-e_0)/2$ applied. If the rectifiers allow leakage current to flow, part of this voltage will be utilized in

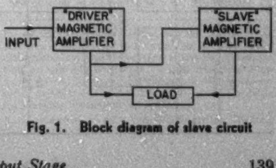


Fig. 1. Block diagram of slave circuit

Paper 55-58, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 15, 1954; made available for printing December 10, 1954.

P. W. BARNHART is with the Feedback Controls, Inc., Alexandria, Va.

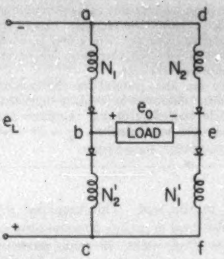


Fig. 2. Half-wave bridge load circuit during reset half-cycle

changing flux in the cores, and control will be established to produce output the next half-cycle independent of any other signal applied to other windings. It will be noted that this control will produce on the next half-cycle an output current in such a direction as to reverse the polarity on the load. If the load polarity, as shown in Fig. 2, were produced by a half-wave bridge magnetic amplifier connected at *b* and *e*, the circuit of Fig. 2 would be a slave stage producing an alternating current with no other control signal. Care must be taken that the roles of the slave and driver stage are not reversed on the next half-cycle. The driver stage must have rectifiers with very high impedances so that feedback from the slave stage will not control it and increase the response times. To produce an alternating voltage across the load with no direct component, it is necessary that the back leakage of the slave stage rectifiers be reduced until the control injected will fire its cores exactly 180 degrees after the respective cores of the driver stage.

Experimental results with a full-wave bridge amplifier indicate that it is impractical to use this method of control. To get the required amount of leakage to eliminate the d-c component, it was necessary to shunt the rectifiers or to remove plates to such a level that the reset windings had to be reversed to maintain equal quiescent currents in both stages. Under such conditions, the slave stage gain was nonlinear and the desired zero direct component could only be obtained at one level. No further work was done at this point to determine whether different combinations of load or reset turns would produce satisfactory output stage with the slave stage controlled in the load windings. Instead, work was shifted to investigation of control through a different winding.

The Use of Separate Control Windings

Possibly the next approach would appear to be adding a control winding to the slave stage and connecting it directly across the load. This method of control was not used, however, because earlier experience had shown that closed loops in half-wave circuits where direct current may flow will, in most cases, decrease the speed of response markedly. Instead, the control windings were connected in series with the load.

This type of connection is peculiarly adapted to the control of a half-wave stage. It is known, and has been proved analytically,[4,5] that the voltage gain of a half-wave amplifier increases as the number of control turns decreases. To achieve this high gain by lowering the control turns, the control source must have a low driving impedance. It must be able to supply the high currents which will flow when maintaining a flux-changing voltage across a low number of control windings. The driver stage designed to drive the motor load can supply these high currents easily. Since only a small voltage is required across the load windings to obtain an output from this high gain slave stage equal to the driver stage output, very little of the load voltage is lost. At the same time, voltages induced into the slave stage control windings under signal conditions are extremely small. This unique combination of properties produces an excellent method of slave stage control. While this type of control could be used in conjunction with the type previously discussed, its use alone is so straightforward and comparatively simple that the other method is probably only of academic interest.

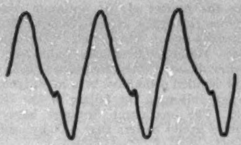


Fig. 4. Full voltage across motor driven by full-wave bridge slave circuit

Fig. 3 shows a full-wave bridge amplifier with control windings on the slave stage; only the load and control windings are shown. Cores 1 and 2 are in the driver stage. It will be noted that the sense of the slave stage control windings is very important. As shown in Fig. 3, the circuit will produce a phase-reversible a-c output. If these windings are reversed, the output will become polarity-reversible direct current. Thus, just as the half-wave bridge, this circuit can be used in conjunction with either a-c or d-c loads. If this amplifier is used to supply direct current, care must be taken that the rectifiers used have a high reverse impedance to prevent control through the load windings which would produce an a-c component.

Using the circuit of Fig. 3 connected for an a-c load, it is possible to apply somewhat more than rated voltage to commonly used 400-cycle-per-second 2-phase servomotors with the control winding in parallel (57.5-volt connection).

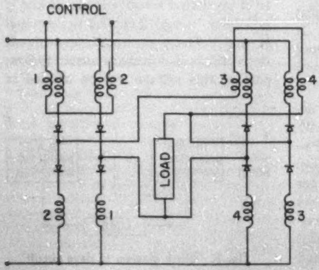


Fig. 3. (left). Full-wave bridge slave circuit for a-c load

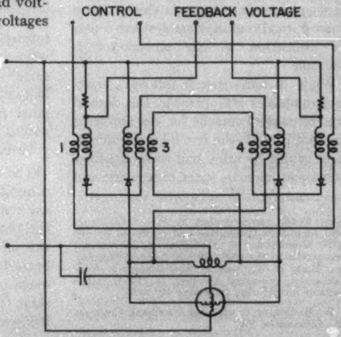


Fig. 5 (right). Full-wave slave circuit with output transformer

The full output wave form across the motor is shown in Fig. 4. The voltage gain of the output stage is about 150.

One disadvantage of the bridge circuit operated directly on a line is that it will not develop full-line voltage across the load. Until recently, it has been somewhat difficult to obtain standard 2-phase servomotors with control windings that could be paralleled for 57.5-volt operation, especially at 60 cycles. With a full-wave bridge it is possible to use a small transformer to increase the motor voltage (even high enough for motors designed for plate-to-plate operation) without danger of saturating the transformer. By center-tapping the transformer, the bridge windings can be simplified to single windings and a full-wave push-pull amplifier with differential output is obtained.

Such an output circuit is shown in Fig. 5. This output transformer can be a simple center-tapped autotransformer if the motor has a 115-volt control winding. It may be tapped in a different manner for motors with higher or lower voltage control windings. If the motor itself has two 57.5-windings or a center-tapped 115-volt winding, it may be used in the circuit of Fig. 5 in the place of the transformer. It is necessary, in this case, to put a capacity across the motor windings to obtain full output. Also the quiescent currents must be set at a level that will not cause excess temperature rise in the motor at zero signal. The slave stage control windings of Fig. 5 are differentially connected. This gives improved linearity over the use of individually controlled cores, the latter case producing interference with reset windings. The number of these control windings used on both 400- and 60-cycle units has been less than 10. The series resistors in the driver stage are used to provide a direct voltage for feedback compensation purposes.[2]

Test Results and Conclusions

The performance of this amplifier in a servo system has been very good. Two half-wave bridge amplifier stages of the type shown in Fig. 2 are used as a preamplifier for the amplifier of Fig. 5 with a 400-cycle motor load (*BuOrd Mk. 8*). With a 10,000-ohm input impedance the open-loop voltage gain is over 700. When the loop is closed with a 50 to 1 gear ratio and using a feedback network providing simple lead compensation, the resulting bandwidth is 17 cycles per second and the velocity constant is 200 degrees per second per degree error. The natural frequency of the mechanical system without compensation is only about 3.5 cycles per second. Lead-lag compensation as used with half-wave amplifiers can also be easily applied.[6]

This circuit has proved to be an excellent output stage for high-performance control amplifiers by combining the advantages of half-wave and full-wave circuitry. However, the use of this type of slave stage control is not limited to driving a-c motors. Controlling phase-sensitive saturable reactors, d-c servomotors, and rotating electromechanical amplifiers are only a few of the other applications which are immediately brought to mind.

References

1. An Improved Magnetic Servo Amplifier, C. W. Lufcy, A. E. Schmid, P. W. Barnhart. *AIEE Transactions*, vol. 71, pt. I, 1952, pp. 281–89.

2. Compensation of a Magnetic Amplifier Servo System, H. H. Woodson, A. E. Schmid, C. V. Thrower. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 8, 1952, pp. 158–65

3. Magnetic Amplifier Circuits With Full-Wave Output and Half-Wave Control Signals. H. W. Lord. *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 265–70.

4. Design Considerations of the Half-Wave Bridge Magnetic Amplifier, C. W. Lufcy, H. H. Woodson. *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 220–25.

5. Preliminary Analysis of the Half-Wave Bridge Magnetic Amplifier, H. H. Woodson. *NavOrd Report 3596*, U. S. Naval Ordnance Laboratory, Washington, D. C., 1954.

6. Lead-Lag and Lead-Integral Servo Compensation Using Half-Wave Bridge Magnetic Amplifiers, H. H. Woodson, L. S. Weinstein, J. E. Roberts. *NavOrd Report 3560*, U. S. Naval Ordnance Laboratory, Washington, D. C., 1954.

Switching Transistors Used as a Substitute for Mechanical Low-Level Choppers

A. P. KRUPER
NONMEMBER AIEE

ONE of the interesting possibilities of using transistors as switches would be to replace mechanical vibrators or choppers in the input of d-c amplifiers. The ability of transistor switches to function with emitter-collector voltages of the order of 1 millivolt or less led to a further investigation of transistor-switching behavior at these very low levels. Peculiarities of circuits utilizing transistors as switches at these low levels are discussed and possible methods of compensating for these peculiarities are presented. Effects of temperature on behavior were also investigated and it was found that, by suitably matching pairs of transistors, the effects of temperature from −50 to +90 degrees centigrade (C) were practically negligible for germanium transistors.

Present Choppers

The chopper applications to be discussed are essentially of the type illustrated in Fig. 1. Here the problem is to amplify a low-level d-c signal in a manner which is free of the drift and instability of d-c amplifiers. A mechanical SPDT switch alternately connects the d-c signal to either side of the center-tapped input transformer producing a square-wave a-c output signal which is easily amplified by the a-c amplifier. A-c amplifiers having good linearity and stability are much easier to design than d-c amplifiers. The output of the a-c amplifier is then rectified, e.g., by a synchronous rectifier and read on the d-c meter. The input chopper can be a motor-driven switch or, what is more common today, a magnetically driven vibrator switch.

In another version of this same circuit, which is perhaps more common for commercial and military applications, the d-c input signal is derived from an error signal which is reduced to a minimum or zero by electric means actuated by an a-c motor driven directly by the output of the a-c amplifier. The direction of rotation of this motor is determined by the phase of

---

Paper 55-137, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing December 3, 1954.

A. P. Kruper is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

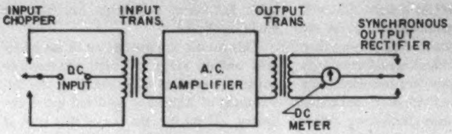


Fig. 1. Schematic representation of chopper-type d-c amplifier

the a-c output signal. It will be observed that, if the polarity of the input d-c error signal is reversed, the a-c signal produced undergoes a 180-degree phase reversal.

Transistor Choppers

The remainder of this paper will concern the properties of this input choppers. Fig. 2 shows a schematic of an electromagnetically driven mechanical vibrator chopper and a direct substitution of the vibrator switch contacts with switching transistors. The mechanical chopper is energized from the line voltage. The switch arm alternately swings from one contact to the other in phase relationship with the line voltage. In Fig. 2(B) an arrangement is illustrated where each switch contact of Fig. 2(A) has been replaced by a transistor switch and each transistor switch is opened or closed by means of an a-c signal applied between the base and collector. However, if the output of the transistor chopper were to be examined, it would be observed that, although it behaved very much like the mechanical chopper, with zero d-c input signal, an a-c signal was still produced in the output. The a-c signal is produced by a small voltage developed by the transistor, of the order of 1 or 2 millivolts, when the switch is closed. This might be likened to the condition which would exist in the vibrator if the switch contacts were of dissimilar metals such that a thermal electromotive force would be produced at the contact points every time the switch was closed. Of course, vibrator manufacturers take great pains to keep these contact electromotive forces to a minimum. The source of this potential in the transistor chopper and possible methods of compensating for it will be investigated in the following sections.

Transistor Characteristics

Fig. 3 shows the family of characteristic curves of a transistor when operated as a switch. As pointed out by Dr. Bright,[1] the transistor does not behave as a perfect switch, a fact which is shown by the displacement of the point *P* from the origin. The voltage appearing across the load of the transistor chopper in Fig. 2 with zero input signal is approximately V_P. The operation of the transistor chopper may be analyzed by considering the load line as the impedance of one-half of the transformer input winding and the d-c input signal voltage as the supply voltage. At this point only the operation of one of the transistors will be considered. As the input d-c signal voltage is decreased, the load line moves downwards towards the origin causing the a-c output voltage to decrease in amplitude. Finally, at some input voltage slightly greater than V_P, the load line passes through *P* and no alternating voltage is produced in the load. Now if the input d-c signal voltage is further decreased until it becomes zero, the load line passes through the origin *O* and an alternating voltage is again developed across the load. The amplitude of this alternating voltage is approximately equal to V_P but is reversed in phase 180 degrees from the alternating voltage which was produced when the load line was above *P*. If the polarity of the d-c input is now changed and increased in magnitude, the load line continues to move downward producing an a-c signal of increasing amplitude. Hence, it is seen that an a-c signal is obtained whose amplitude is proportional to the magnitude of the d-c input and whose phase is a function of the polarity of the d-c input.

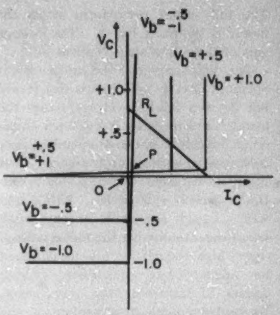


Fig. 3. Characteristics of n-p-n switching transistor at low levels

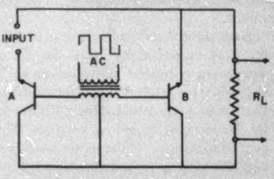


Fig. 4. Improved transistor chopper circuit

A word might be mentioned at this point about the order of magnitude of the coordinates V_P and I_P for the normal and inverted connections. Values between 10 and 40 millivolts have been obtained for the normal connection of typical n-p-n units, while values for the inverted connection for the same transistors ranged from 0.5 to 1.5 millivolts. Currents of the order of 40 microamperes in the normal connection were reduced to 1 microampere in the inverted connection at room temperature. This clearly points out the superiority of the inverted connection for low-level chopper applications.

Circuit Considerations

Whereas the transistor chopper illustrated in Fig. 2 suffered from the effects of this residual voltage V_P, or feed-

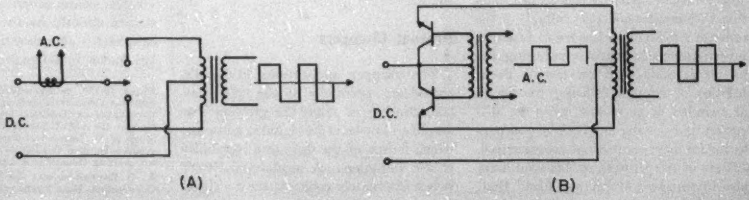


Fig. 2. Input chopper circuits showing (A) conventional mechanical vibrator, and (B) substitution of switching transistors for mechanical contacts

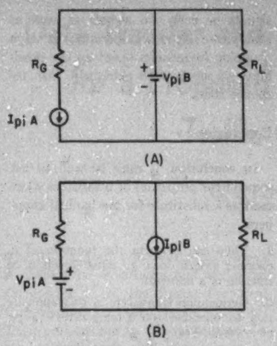


Fig. 5. Equivalent circuits of transistor chopper shown in Fig. 4

through voltage as it is sometimes called, Fig. 4 shows a circuit in which these voltages from the two transistors tend to cancel each other providing a nearly ideal low-level chopper. The input is applied as shown and, on one half-cycle of the chopping frequency, transistor *B* is open and transistor *A* is closed, completing the circuit from the input through *A* and to the load R_L. On the alternate half-cycle, transistor *A* is open and transistor *B* is closed, effectively short-circuiting the load R_L. Transistors *A* and *B* are shown in the inverted connection and are preferably driven by a square wave of approximately 1 volt of amplitude.

In this particular circuit, the maximum input voltage is limited to the same order of magnitude as the base driving voltage if the transistors are to operate satisfactorily as switches. This is ordinarily no objection in low-level operation where the input signal always remains in the millivolt region such as when operating with thermocouple input. If it is desired to work with higher input voltages and still to go down into the low-level regions, then the back-to-back transistor arrangement mentioned in reference 1 can be used in place of *A* and *B*, which would then require four transistors instead of two. Such a circuit could be used to convert voltages of the order of 50 volts of direct current of either polarity to alternating current, such as is required for analogue computers. Extremely good linearity of conversion to a fraction of a per cent can be achieved in this application.

Fig. 5 shows the equivalent circuit for the low-level chopper during alternate half-cycles. During the first half of the cycle, as shown in Fig. 5(A), transistor *B* is conducting and is represented by a battery V_{PiB} of the polarity shown directly across the load, and the transistor *A* is open and is represented by a current generator I_{PiA} in series with the input source R_G. Here the voltage developed across the load is the voltage caused by transistor *B*, and the effect of the current from *A* is negligible. During the alternate half-cycle, transistor *A* is closed and is now represented by a battery in series with the input source R_G. Transistor *B* is open and is shown as a current source across the load R_L. If R_G is small, such as would be the case with a thermocouple input, the voltage across the load is essentially the input voltage in series with V_{PiA}, and the effect of I_{PiB} is negligible. For the condition of no d-c input signal voltage, the voltages from transistors *A* and *B* appear across the load during alternate half-cycles but in the same direction and, hence, produce a constant direct voltage rather than an alternating voltage across the load, if they were initially selected to have equal values of V_P.

The effect of I_{Pi} was stated as being negligible due to the fact that its magnitude was small. This might not be the case for germanium transistors at high temperatures and, hence, R_G should be kept as small as possible. Alternately, it is possible to insert a series resistance equal to R_G in the leg of transistor *B* such that the voltage drops caused by the I_{Pi}'s are equal. In this manner the equivalent circuits for both halves of the cycle become identical providing nearly perfect compensation.

Temperature Effects

Since the problem of compensating for the feed-through voltage V_{Pi} by matching transistors has been mentioned, it might be of interest to examine some typical values obtained for V_{Pi} for six unselected experimental Westinghouse n-p-n germanium transistors. Fig. 6 shows a plot of V_{Pi} as a function of temperature for these transistors. It will be seen that the maximum spread in this value is about 1 millivolt over a temperature range of −50 to +90 C. Also, it will be noted that four of the six units tested all fell within a region where the spread was about 0.1 millivolt over this same temperature range. It would not be difficult to select two transistors, e.g., numbers 1 and 5 which would match to within about 25 microvolts of each other over this temperature range. Thus, it is

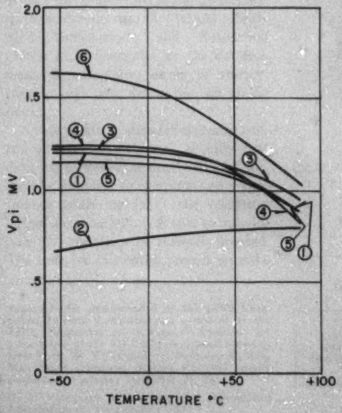
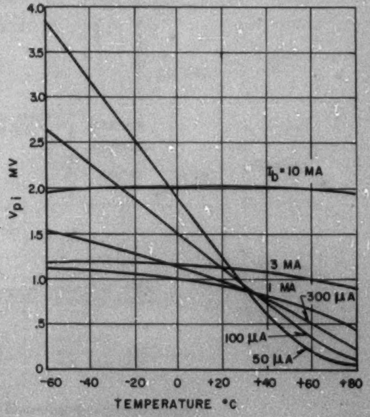


Fig. 6 (left). Variation of V_{Pi} as a function of temperature for six randomly selected n-p-n junction transistors

Fig. 7 (right). Variation of V_{Pi} of a typical n-p-n transistor as a function of temperature and base drive



seen that it would be possible to use these two transistors in a transistor chopper circuit and have less than 25 microvolts' a-c signal output produced over a temperature range of −50 to +90 C if zero-input signal voltage were applied.

Similar measurements on the normal connection for the same six transistors exhibited a spread of from 10 to 40 millivolts of the feed-through voltage V_{Pn} and the best compensation which might be achieved over the same temperature range would be about 2 millivolts at best. This, again, points out the superior advantages of the inverted connection.

The curves in Fig. 6 show the variation of V_{Pt} with temperature with a constant base drive of about 3 milliamperes. Fig. 7 shows a plot of the variation of V_{Pt} with temperature as a function of the base drive current for a typical transistor. It should be noted that, as the base current is increased, the curves become more nearly horizontal. This provides a convenient means of matching the feed-through characteristics of pairs of transistors, since a series resistor in the base circuit of each transistor can be adjusted to provide the desired base current. Alternately, since the curves are so nearly horizontal, it is possible in certain applications to buck out this voltage with a fixed potential.

Advantages of Transistor Choppers

Now that the behavior of a transistor chopper circuit has been reviewed and it appears that it is possible to operate such a device at levels of 1/10th millivolt and less over considerable temperature variations, other advantages of transistors for this application should be mentioned.

One of their big advantages over mechanical choppers is that they are inertialess devices. As such, they can be operated at any line frequency, or even varying frequency, without loss in performance. They have practically zero phase shift up to about 10 kc as contrasted to appreciable phase shifts from 30 to as high as 90 degrees for mechanical choppers. Mechanical choppers, on the other hand, are usually designed for a specific frequency or narrow range of frequencies because of problems associated with designing the mechanical vibrating reed, and the phase shift also varies with chopping frequency.

Being inertialess devices of very small mass, transistors are practically immune to shock and vibration and, hence, are ideal choppers for military and guided missile applications. Actual tests under all environmental and temperature conditions have not been completed but are in the process of being investigated.

Although nothing has been said about other than mechanical choppers, it should be pointed out that other modulators such as diode-bridge modulators, etc., can be replaced by the transistor choppers with corresponding gain in sensitivity at low levels.

The fact that the transistor chopper can be operated at frequencies up to hundreds of kilocycles makes it possible to use it as a carrier for rapidly varying d-c signals or even a-c signals as well as steady d-c signals. Operation at high chopping frequencies opens up the possibility of servos with extremely fast response times.

Conclusion

In conclusion, it may be well to list some of the properties of transistors when used as a substitute for mechanical choppers.

1. They can perform the function of a chopper switch down to signal levels of a fraction of a millivolt.
2. Germanium transistors in this application are operable over a temperature range of −50 to +90 C.
3. They are synchronous with line frequency and have practically zero phase shift up to the kilocycle ranges.
4. They are low-impedance devices and should, therefore, be used with low-impedance sources.
5. They are small in size and light in weight.
6. They are nonmicrophonic and insensitive to shock.
7. They should have inherently long life as contrasted to mechanical choppers.
8. They are not subject to contact pitting and burning as occurs in mechanical choppers.
9. They have a dynamic range of greater than 110 decibels.

Reference

1. JUNCTION TRANSISTORS USED AS SWITCHES, R. L. Bright. *AIEE Transactions*, vol. 74, pt. I, Mar. 1955, pp. 111–21.

Vol
Pt,
195

Data Transfer and Display Equipment for a Proposed System of Air Traffic Control

G. E. FENIMORE
NONMEMBER AIEE

IN 1947, a group representing all phases of the aviation industry met as Special Committee 31 of the Radio Technical Commission for Aeronautics (RTCA). Their purpose was to bring order out of a chaos where many different systems of air navigation and air traffic control were being proposed and developed to satisfy the same need. This group evaluated the existing and proposed systems and formulated a long-range plan for orderly progress toward the development of a common system to serve all nontactical military air traffic and all civilian aircraft.

The *SC-31* report which the group produced outlined the broad general requirements for an automatic air traffic control system using information transmitted over land line teletypewriter circuits and ground-to-air radio circuits.[1] The use of several special-purpose computers to process the data acquired was also recommended in the report.

In 1948, the Air Navigation Development Board was established to administer funds which were appropriated by Congress to develop the common system recommended by the RTCA. A study was made of the use of computers for handling air traffic data, and in 1950 a project was established at the Civil Aeronautics Administration (CAA) Technical Development and Evaluation Center in Indianapolis for the "development of automatic means of storing, transferring, and displaying air traffic data.

An automatic system for air traffic control on a nation-wide scale requires that data for numerous remotely located aircraft be supplied to computers on a continuous basis and that the computer output be promptly available to individuals who may also be remotely located. The time for this entire process must be short enough, so that air traffic controllers concerned with a predicted confliction between aircraft can be aware of the difficulty in time to reassign airways or altitudes to avoid an accident. Such a system could be likened to a computer having its input and output devices remotely located from the memory and computation equipment.

It was determined that basic building blocks for the system could be represented by a fully automatic teletypewriter switching system for communications, magnetic-drum storage equipment for memory and processing of flight plan information, and automatic display boards for rapid data display and transfer. The proposed organization of this system is shown in Fig. 1, which illustrates how the equipment may be used for handling both flight plan data and weather information.

A flight plan includes such information as aircraft identity, departure airport, time of departure, destination, cruising speed, and a proposed timetable for the flight. Data in the form of incoming flight plans may be fed into the system from a remote station connected to the automatic teletypewriter switching center. Such remote stations would be located at major airports and at CAA communication stations where flight plans are normally filed. This teletypewriter message would be directed by the automatic switching center to the magnetic drum equipment for storage and would be displayed in front of an air traffic controller for approval at the local display board. The traffic controller would check the proposed flight plan against other flights for conflicts and would issue a clearance for the flight. In the first implementation of the system, this check would be made by the traffic controller, but an ultimate system may include a computer to search for conflicts automatically, using data stored on the drum. Clearances would be automatically directed to the remote display at the departure airport tower.

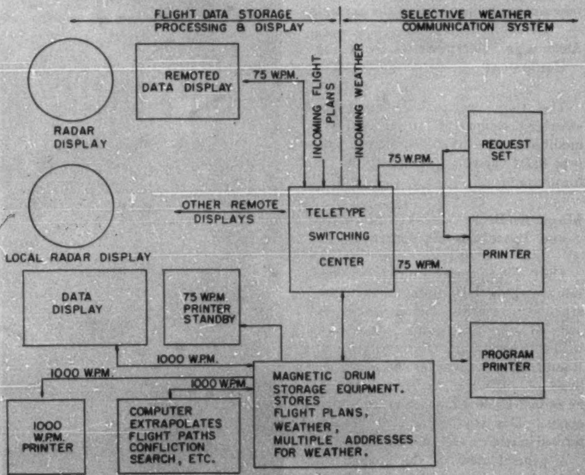

FLIGHT DATA STORAGE PROCESSING & DISPLAY
SELECTIVE WEATHER COMMUNICATION SYSTEM
RADAR DISPLAY
REMOTED DATA DISPLAY
75 W.P.M.
INCOMING FLIGHT PLANS
INCOMING WEATHER
REQUEST SET
75 W.P.M.
LOCAL RADAR DISPLAY
OTHER REMOTE DISPLAYS
TELETYPE SWITCHING CENTER
PRINTER
75 W.P.M.
DATA DISPLAY
75 W.P.M. PRINTER STANDBY
PROGRAM PRINTER
1000 W.P.M.
1000 W.P.M.
1000 W.P.M.
MAGNETIC DRUM STORAGE EQUIPMENT. STORES FLIGHT PLANS, WEATHER, MULTIPLE ADDRESSES FOR WEATHER.
1000 W.P.M. PRINTER
COMPUTER EXTRAPOLATES FLIGHT PATHS CONFLICTION SEARCH, ETC.

Fig. 1. Block diagram of proposed system

While the flight is in progress, information must be transferred from one traffic control center to another along the route. Controllers are in contact with the flight by radio and may need to alter schedules or clearances from time to time because of delays or new flights entering the area. Drum information must be kept current and traffic controllers must have current schedules available at all times. These

---

Paper 55-148, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 22, 1954; made available for printing December 21, 1954.

G. E. FENIMORE is with the Civil Aeronautics Administration, Indianapolis, Ind.

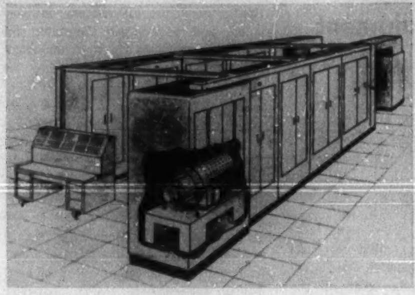


Fig. 2. General view, magnetic-drum storage equipment

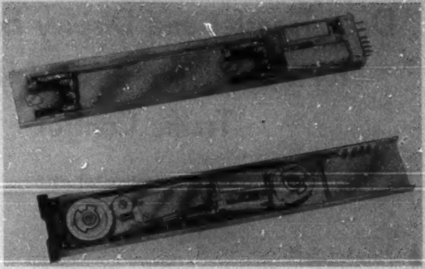


Fig. 4. Inside view of indicator

data would be provided by the proposed automatic system.

On the right-hand side of the diagram, the same teletype switching center and magnetic-drum storage equipment are involved in a proposed system of collecting, storing, and distributing aviation weather information. Incoming weather reports such as the Cleveland weather, for example, pass via the switching center to the drum where they are stored and automatically sent to the program printer at locations requiring this weather information on a continuous basis. Stations which need the Cleveland weather occasionally would be able to send a message to the drum requesting this information and to receive in reply the desired weather report. Having discussed the proposed system operation, it will be necessary to take a closer look at some of the elements.

Automatic Teletypewriter Switching Equipment

The fully automatic teletypewriter switching equipment to be used is a modified version of the Western Electric type-*81D1* system, described in the literature.[2,3]

Magnetic-Drum Message Storage and Processing Equipment

The use of magnetic drums for information storage is well known in the computer art. The magnetic-drum message storage and processing equipment associated with this system was built by Remington Rand Inc., and is now operating on a test basis at the CAA Technical Development and Evaluation Center in Indianapolis. This equipment has been described in references 4 through 6.

A general view of the storage and processing equipment is shown in Fig. 2. Fig. 3 illustrates the magnetic drum which is capable of storing 2,000 flight plans in the form of individual teletypewriter messages. The total information storage capacity of the drum is 316,250 characters or 1,581,250 bits. When a request for information has been received by this equipment, the time required for searching through the drum storage and preparing the desired message for replay is less than 1/2 second.

Data Transfer and Display Equipment

The third element of this information processing system to be described in detail is the data transfer and display equipment. This equipment has been developed to satisfy the following basic requirements:

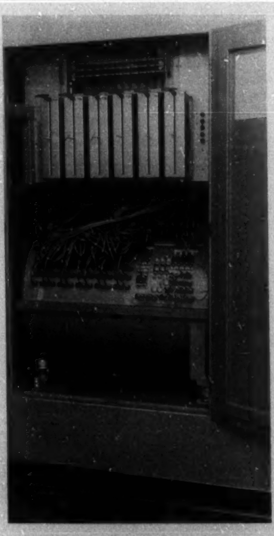


Fig. 3. Magnetic-drum cabinet

1. Display of information: To display on a board in front of air traffic controllers information such as the identity of an aircraft, its altitude, speed, destination, and a timetable of its route. This information constitutes the flight plan.
2. Ability to communicate with the magnetic-drum equipment: To permit an operator to send a request message to the magnetic-drum storage equipment asking for the flight plan of any aircraft and display automatically the reply from the drum on his board.
3. Ability to change individual characters or the sequence of display information: To permit the operator by simple push-button manipulation to rearrange the order in which various flight plans appear on his board or to change individual characters of the displayed information.
4. Rapid data transfer between display boards: To permit the operator by simply push-button manipulation to transfer information appearing on his board to another similar board, or to compose and send automatically a message containing the displayed information to the magnetic-drum equipment.

Specifications for equipment to satisfy these requirements, and others which will be described later, were written by the Technical Development and Evaluation Center of the CAA. The Union Switch and Signal Division of Westinghouse Air Brake Company, has developed and produced two display boards for evaluation in air traffic control systems.

Indicators

The basic building block of the display boards is a small electromechanical indicator which can display any one of 64 different letters, figures, or symbols in a small window at one end. Each indicator is approximately 11 inches long,

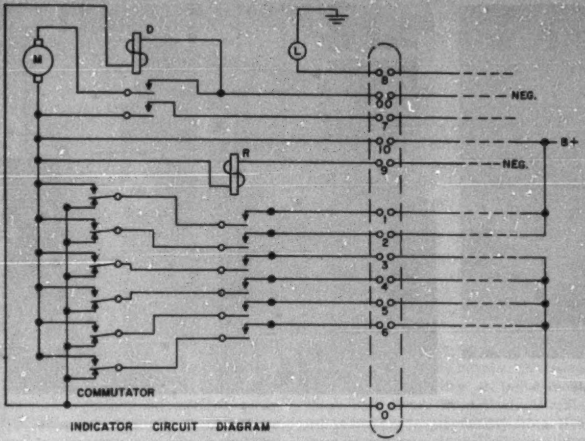


Fig. 5. Indicator circuit diagram

$1^3/_4$ inches high, and 5/8 inches wide and is capable of changing the character displayed to any other character within $1^1/_2$ seconds. A number of such indicators can be mounted in a console and a single row can be energized to spell out information. Since the magnetic-drum storage equipment and the automatic switching center both handle information in the form of teletypewriter code, the indicators and display equipment have been designed to speak the common language.

Fig. 4 is an inside view of one of these indicators. At the rear is a moulded plug for making the electric connections. These connections include six code wires for determining which character shall be positioned, and several wires for power, lighting, and control. Near the plug is a relay which, when energized, will make connections to the six code wires. At the center is a small motor which drives the sprocket and endless belt assembly during positioning. A detent magnet acting upon the rear sprocket insures accurate positioning of each character in the front window and controls power to the motor. On each sprocket wheel is a code segment which together with the brushes shown on the upper half of the indicator provide a homing circuit with the six code wires for positioning a desired character.

An electrical diagram of the indicator is shown in Fig. 5. For a character such as the letter *A* to be positioned, a direct voltage must be applied to wires 1 and 2 while wires 3, 4, 5, and 6 are connected together, as shown in the diagram. This corresponds to the teletypewriter code for *A* which is two marks and three spaces in the 5-unit system. Absence of voltage on the sixth wire indicates that this is a lower case character. Since a 6-wire system can provide 64 combinations, the indicator is capable of displaying 64 unique characters in its window. The characters selected for this model are 26 letters, 10 digits, and the weather symbols commonly used in the CAA teletypewriter circuits.

In addition to the ability to display various characters, the indicator acts as a memory device in that it is also capable of reading out the electrical code for the character being displayed. By energizing wires 9 and 10, this read-out function can be accomplished without disturbing the displayed information. Under this condition if the indicator is at rest with the character *A* being displayed, wires 1 and 2 will have voltage on them, while 3, 4, 5, and 6 will be connected together. With two indicators connected back to back, one indicator can be energized to read out its character and the other indicator will be driven until it is displaying the same character as the first. This principle is utilized for transferring information from one row of indicators to another in the console.

Manual Operation of the Display

The display console is shown in Fig. 6. The sloping panel can accommodate 660 indicators arranged in 11 rows of 60 each. Data concerning one flight plan normally would be displayed in one row. Certain buttons and lamps are associated with each individual row and these are arranged in groups along the left side of the indicators, see Fig. 7. These buttons provide normal control for the following operations:

1. Data in any row can be transferred to any vacant row.
2. Data can be erased from any row.

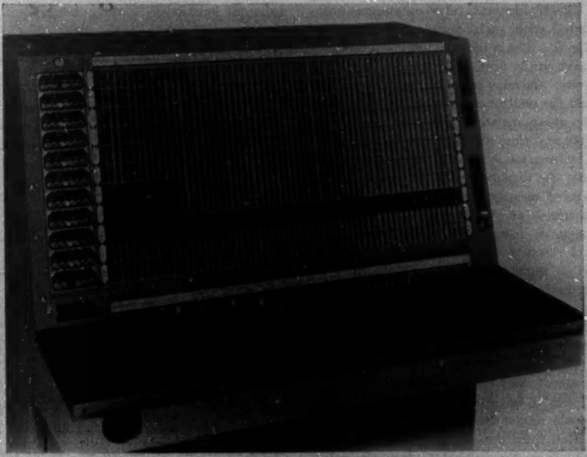


Fig. 6. Display console

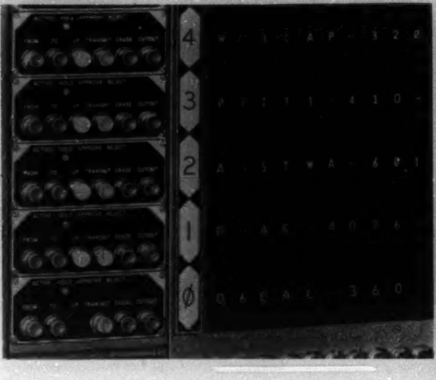


Fig. 7. Close-up view of console

Fig. 8. Display equipment racks

3. By operation of a single button, all data can be concentrated in the lower rows of the display console without changing the prevailing sequence.

4. Similarly, by operation of only one button, a selected row can be cleared of information by transferring the data upward.

5. Characters displayed in a row of indicators can be altered individually or in groups.

6. Data appearing in a row can be composed automatically into a message and transmitted at any speed up to 1,000 words per minute.

7. Individual rows can be removed from service.

Electronic and relay equipment associated with each console are shown in Fig. 8. The apparatus mounted on these racks can be divided into three functional units: the relay register, the shift register, and the control circuits. The purpose of the relay register is to store one flight plan at a time so that it can be processed. The function of the shift register is to enable the display system to transmit and receive messages at any speed and it is designed to be compatible with the magnetic-drum equipment. The control circuits sequence the display, the relay register, and the shift register through their various functions, check for proper operation, detect errors, and malfunctions, and provide alarms to the maintainer or operator.

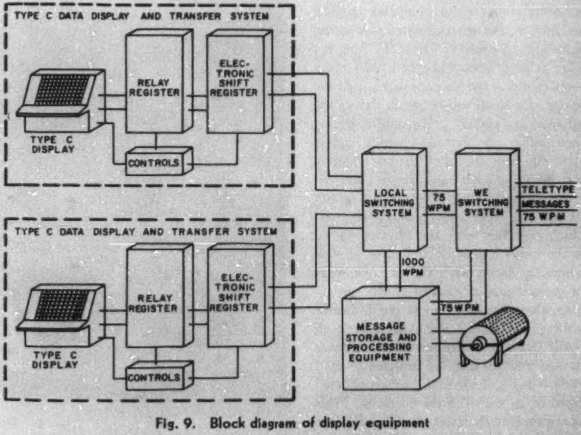
Fig. 9 is another block diagram of the system equipment with the displays shown in greater detail. Data which appear on the console board can be transmitted from the display by first moving data from the indicators into the relay register, transferring to an electronic shift register, and feeding the information into the transmission line where it will be automatically switched to the proper destination. Incoming messages intended for the display equipment are first received by the electronic shift register as teletypewriter characters, transferred to the relay register, and then to indicators in the display.

The display equipment is designed to handle any communication speed up to 1,000 words per minute. The normal speed will be 1,000 words per minute when communicating with a local drum or another display and 75 words per minute when the message is destined for a distant location.

Automatic Operation of the Display

In normal operation, the display boards are in readiness to receive incoming messages. Such messages, when received, are sensed for a "kind" character which occurs in the preamble of the message. This kind character serves to identify the purpose of the message and initiates an internal program control. The program control can cause one or more of the following actions to be performed:

1. The message may be displayed in an empty row.



TYPE C DATA DISPLAY AND TRANSFER SYSTEM
RELAY REGISTER
ELECTRONIC SHIFT REGISTER
TYPE C DISPLAY
CONTROLS
TYPE C DATA DISPLAY AND TRANSFER SYSTEM
RELAY REGISTER
ELECTRONIC SHIFT REGISTER
TYPE C DISPLAY
CONTROLS
LOCAL SWITCHING SYSTEM
75 WPM
WE SWITCHING SYSTEM
TELETYPE MESSAGES 75 WPM
1000 WPM
75 WPM
MESSAGE STORAGE AND PROCESSING EQUIPMENT

Fig. 9. Block diagram of display equipment

2. The rows may be searched to find a displayed flight number which is identical to that contained in the incoming message. The message may then be read into the proper row.

3. Colored lamps may be lighted or flashed to indicate approval or rejection of flight clearances.

4. A reply message may be automatically composed and transmitted.

5. A cancel message may automatically erase information displayed in a particular row; this would normally be the result of action initiated by the operator himself.

Numerous other automatic operations can be performed by the equipment, both in transmission and reception of information, to provide error checking and to prevent loss of information by improper button operation. By combining push-button operation and automatic program operation, the display boards have been designed to perform automatically many functions which air traffic controllers are now handling manually, thus permitting the controller to devote more time to important decisions which should improve the efficiency and safety of air traffic.

Summary

The density of air traffic has increased to the point that the present traffic control system finds it difficult to handle peak loads without causing some delays in schedules. The air traffic controller is required to make numerous manual operations and to scan visually a large amount of air traffic data in performing his duties. An automatic system is being developed wherein all air traffic information, by being reduced to electrically coded form, can be handled by equipment such as automatic switching centers, magnetic-drum storage and processing, and automatic display boards. A testing program is under way at Indianapolis where both simulated and actual traffic will be handled by present-day means and by new systems. It is likely that the best system of air traffic control will combine new operational procedures with new tools such as improved radar coverage and automatic information handling equipment.

References

1. Air Traffic Control, **Special Committee 31.** *Paper 27–48/DO-12,* Radio Technical Commission for Aeronautics, Washington, D. C., May 12, 1948.

2. A Full Automatic Private-Line Teletypewriter Switching System, **W. M. Bacon, G. A. Locke.** *AIEE Transactions,* vol. 70, pt. I, 1951, pp. 473–80.

3. The 81-C-1 Teletypewriter Switching System, **Walter M. Bacon.** *Bell Laboratories Record,* New York, N. Y., vol. 28, no. 4, April 1950, pp. 145–52.

4. Message Storage and Processing with a Magnetic Drum System. **A. P. Hendrickson, G. I. Williams, J. L. Hill.** *Proceedings,* Eastern Joint Computer Conference, Philadelphia, Pa., Dec. 1954.

5. Applicability of Magnetic-Drum Information Storage to the CAA Teletypewriter Circuits, **Francis J. Gross.** *Report No. 233,* Civil Aeronautics Administration, Indianapolis, Ind., April 1954.

6. An Electronic System for Processing Air Traffic Control Information, **R. M. Kalb.** *AIEE Transactions,* vol. 74, pt. I, May 1955, pp. 152–55.

Some Properties of Magnetic Fluids

J. E. COOLIDGE, JR.
ASSOCIATE MEMBER AIEE

R. W. HALBERG
NONMEMBER AIEE

SINCE the original announcements of J. Rabinow's investigations of magnetic fluid clutches in 1948,[1] much energy has been expended to make practical use of this phenomenon. Application to computors, machine tools, dynamometers, and automotive clutches are but a few of the areas which have been studied. The results of these application studies have not been as rewarding as originally anticipated, although at least two manufacturers are making a commercial product.

Several years ago, the Borg-Warner Central Research Laboratory initiated a program of insurance research on magnetic fluid clutches. The work has proceeded along two paths: the measurement of the properties of magnetic fluids; and observation of the characteristics of running clutches. This paper is a report of some of the results of the first study.

Basic Principles

A magnetic fluid clutch takes advantage of the solidification of iron powder and lubricant mixtures when exposed to a magnetic field. This control of the state of the fluid allows variation of the force transfer between the elements confining the fluid and, thus, a variable capacity clutch. Theoretically, the form of the clutch elements is only limited by the means to obtain proper magnetic excitation, but operating characteristics provide further limitations.

Objectives of the Program

The initial aim of the magnetic fluid study was to find the relation of material type, particle density, particle shape, fluid properties, flux density, and gap thickness to the shear-stress or force-transfer capacity of the magnetic fluid. Although the study has left many questions unanswered and avenues unexplored, the partial results are presented in the hope they might stimulate additional work and, thus, provide a better understanding of this phenomenon.

Equipment

All shear-stress measurements were carried out on the apparatus shown in Fig. 1. This fixture consists of a U-shaped electromagnet with round pole faces of 0.44 square inch in area, the centers of which are 1.16 inches from the center of rotation, as shown schematically in Fig. 2. The magnet is mounted on a freely rotatable center shaft with spacers between a shaft collar and the magnet to provide an adjustment in gap between the pole faces and the base from zero to 0.030 inch in 0.002-inch intervals and, in larger intervals, of 0.030-inch up to a 0.060-inch gap.

Directly below and parallel to the magnet pole faces, poles of slightly greater area are machined in the steel base plate. The base poles are encircled by calibrated brass rings which are threaded to enable adjustment of the depth from the surface of the ring to the base pole surface. The purpose of the retaining rings is to insure that a constant mixture concentration is exposed to the uniform magnetic field of the pole pieces. This construction necessarily restricted the movement of the magnet and made measurement of the shear stress difficult. However, after much experimenting, a satisfactory technique was developed.

Flux density measurements were made by a 25-turn search coil and a ballistic galvanometer. The search coil is wound as close to the end of the magnet pole face

---

Paper 55-170, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted May 14, 1954; made available for printing November 24, 1954.

J. E. Coolidge, Jr., and R. W. Halberg are with the Borg-Warner Central Research Laboratory, Bellwood, Ill.

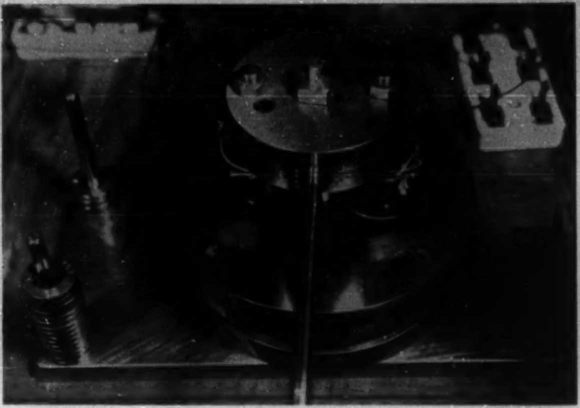


Fig. 1. Assembly, static torque fixture

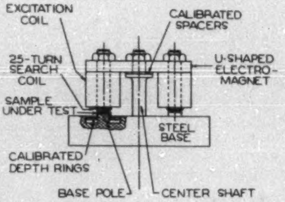

EXCITATION COIL
CALIBRATED SPACERS
25-TURN SEARCH COIL
U-SHAPED ELECTRO-MAGNET
SAMPLE UNDER TEST
STEEL BASE
CALIBRATED DEPTH RINGS
BASE POLE
CENTER SHAFT

Fig. 2. Schematic drawing of magnetic circuit

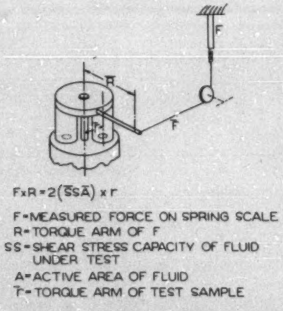

F
R
F
r
FxR=2(SSA) x r
F=MEASURED FORCE ON SPRING SCALE
R=TORQUE ARM OF F
SS=SHEAR STRESS CAPACITY OF FLUID UNDER TEST
A=ACTIVE AREA OF FLUID
r=TORQUE ARM OF TEST SAMPLE

Fig. 3. Schematic drawing of complete apparatus for measuring shear-stress capacity of magnetic fluids

as possible. The shear-stress characteristics of the fluids were obtained by means of torque measurements, using the spring scale and associated gear. The complete setup is shown in Fig. 3. Because of the restricted movement of the magnet, the initial yield stress was used as a comparative stress criterion rather than the maximum or ultimate stress. The yield stress of a mixture can be thought of as the transition point between its action as an elastic body and that of a viscous fluid. The yield stress was determined in the following manner. With zero magnet current, a load was set by holding the magnet in position by hand and the spring tension was adjusted. The electromagnet was then energized and the hand released. The maximum load which produced no movement of the magnet following the hand release have the yield stress.

Test Procedure

After the height of the brass ring had been adjusted to the desired gap and the corresponding spacers put on the magnet shaft, the material to be studied was placed on each base pole. It was leveled by means of a razor blade. The unit was then assembled and torque readings taken at various magnet excitations. A variation of less than 4 per cent was observed for any set of observations.

Results

Extensive observations were made on 14 materials; six Carbonyl and eight hydrogen-reduced irons at flux densities up to 9,000 gauss, working gaps of 0.004 to 0.060 inch, and iron-to-oil mixture ratios of 4 to 1, 6 to 1, and 8 to 1. Although dry mixes could not be studied, some measurements were taken with sus-

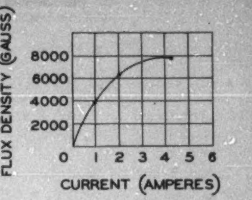


Fig. 4. A typical flux-excitation relation of the magnetic fluids. IRN-6, 6-to-1 ratio

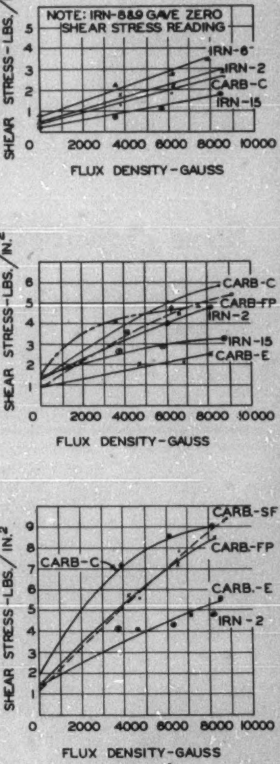


Fig. 5 (top), Fig. 6 (center), and Fig. 7 (bottom). Shear stress versus flux density of representative magnetic fluids of a ratio of iron powder to oil by weight

Fig. 5: 4 to 1 ratio
Fig. 6: 6 to 1 ratio
Fig. 7: 8 to 1 ratio

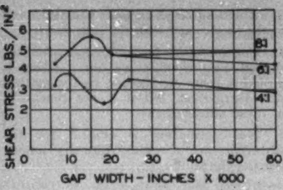


Fig. 8. Shear stress versus gap width for typical hydrogen-reduced iron powder fluids. IRN-2, 8,000 gausses

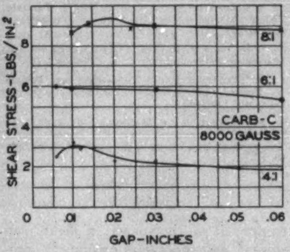


Fig. 9. Shear stress versus gap width for a typical Carbonyl iron powder fluid

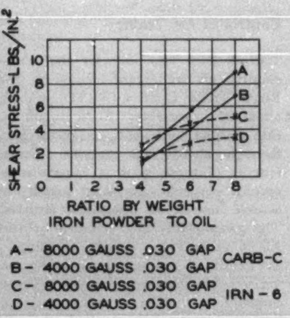


Fig. 10. Shear stress versus iron-powder-to-oil ratio by weight of comparative hydrogen-reduced and Carbonyl iron powder mixture

pensoids of a wide range of initial viscosity. A typical flux-excitation curve is shown in Fig. 4.

Figs. 5, 6, and 7 illustrate the variation of yield stress as a function of flux density for representative irons at 4 to 1, 6 to 1, and 8 to 1 iron-oil mixtures at a 0.030-inch gap. Figs. 8 and 9 show the variation of shear stress with gap thickness for a representative hydrogen-reduced iron and Carbonyl iron at several operating mixtures. Fig. 10 illustrates the variation of shear stress with iron-to-oil ratio at a constant flux excitation and a gap of 0.030 inch for the same irons as shown in Figs. 8 and 9. The physical characteristics of the various powders are given in Table I.

Table I. Characteristics of Powders

Type*	Average Particle Diameter, Microns	Size Range, Microns	Particle Density, Grams Per Centimeter^3	Applied Density, Grams Per Centimeter^3	Particle Shape
IRN-6	8	5 to 20	7.8	2.4 to 2.5	nearly round
IRN-2	15	10 to 30	7.1	2.4 to 2.75	nearly round
IRN-15	15	10 to 30	7.7	1.9 to 2.1	nearly round
IRN-31	40	25 to 150	7.6	3.0 to 3.15	irregular
IRN-8	6	5 to 20	5.1	1.2 to 1.45	nearly round
IRN-9	8	5 to 20	5.1	1.45 to 1.55	nearly round
IRN-3	20	10 to 40	7.9	2.7 to 2.8	nearly round
IRN-16	8	3 to 20	7.8	1.75 to 2.15	nearly round
Carb. *L*	20		7.85	1.8 to 3.0	conglomerated round
Carb. *HP*	10		7.86	2.5 to 3.0	round
Carb. *C*	10		7.86	2.5 to 3.0	round
Carb. *E*	8		7.77	2.5 to 3.5	round
Carb. *TH*	5		7.79	2.5 to 3.5	round
Carb. *SF*	3		7.81	2.5 to 3.5	round
Experimental carbonyl	/1				thistle-shaped
Plast iron					
Sponge iron					flake

* IRN stands for irons and Carb. for Carbonyl.

Discussion

Flux Density

The variation of shear stress with flux density observed in these tests was not as expected. Although other investigators[2,3] have reported a relation B^n, n ranging from 1.6 to 2, a linear relation was most predominant in these measurements. Subsequent tests on running clutches have shown that this difference is probably caused by additional mechanical packing of the magnetic mixture in plate-type clutches. Brede[4] found a linear relation in his drum-type adaption for dynamometer use.

Iron-Oil Mixture and Iron

The results show that the iron-oil ratio and individual irons are both important in determining the shear-stress capacity of the mixture. Actually, the iron-oil ratio is a density effect, as shown by representative irons in Fig. 11. Two significant differences can be observed in this graph: the rate of change of stress with mixture density over this density range among irons of the same type are equal; and the shear-stress capacity of normal carbonyl irons increases much more rapidly than that of hydrogen-reduced irons.

The relation of particle size and shape to these differences is not clear. There definitely appears to be no relation between average particle size and the stress level among the same type over the range 3 to 40 microns. There is some evidence, however, that particle shape has an influence on the slope of the stress-density relation. The characteristics of an irregularly shaped carbonyl iron and an electrolytic iron were found to be similar to the hydrogen reduced irons, as shown in Fig. 11.

The low density IRN-8 and IRN-9 irons failed to indicate any shear-stress capacity at the maximum flux excitation. Extremely low capacities were also observed on a low-density flakey sponge iron and a flakey stainless steel, although the average particle size was much greater.

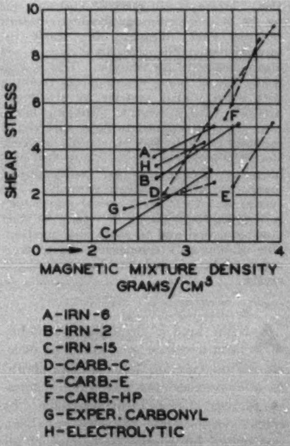


Fig. 11. Shear stress versus magnetic density for several iron powders

GAP THICKNESS

The unusual variation of shear stress with working gap, shown in Figs. 8 and 9, is puzzling. When first observed, it was felt that the structure was something peculiar to the test apparatus but the variation among powders and reproducibility of the measurements dispelled this theory. The Carbonyl irons showed little variation between 0.006 to 0.060 inch with a slight maximum occurring between 0.010- and 0.020-inch gap.

The hydrogen-reduced iron showed a much greater and more erratic variation than the Carbonyls, as high as a 50-per-cent change in stress accompanying a 0.004-inch change in working gap. The significance of these variations is not known.

SUSPENSOID

In all the measurements discussed so far, a light transmission oil was used as a suspensoid. Some observations were made using Aroclor, *DC-200* (silicone), and Hevyteck 95, and little variation in shear-stress capacity was noted. Subsequent measurements in running clutches showed that this independence extended to even dry mixes. Then results concur with those of a previous investigation.[5]

Conclusions

It is believed that the following conclusions can be drawn from this investigation:

1. There is a great difference in the shear-stress capacities of iron-powder fluid mixtures, with particle density being an influential factor. Some of the minor variations may be masked in an actual clutch installation because of mechanical packing and mixture concentration, but the over-all variation is significant.
2. The shear-stress capacity of a magnetic fluid is proportional to the flux density.
3. The shear-stress capacity is independent of the fluid mixed with the iron powder.
4. For maximum force transfer, spherical-shaped iron powders are best. There is probably some significance to the particle size distribution, but it is not conclusive in these measurements.
5. No explanation is evident for the variation in shear-stress capacity among irons of the same type.

References

1. **F. Rabinow.** *Technical Report*, Bureau of Standards, Washington, D. C., 1948.

2. DETERMINATION OF DYNAMIC CHARACTERISTICS OF A MAGNETIC FLUID CLUTCH, **A. J. Parziale.** *Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., June 1949.

3. THE PROPERTIES AND PERFORMANCE OF THE MAGNETIC FLUID CLUTCH, **C. A. Cibelius.** *Ibid.*

4. A MAGNETIC FLUID DYNAMOMETER, **Dwight W. Brede.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1,652–55.

5. INVESTIGATION OF MAGNETIC MIXTURES FOR CLUTCH APPLICATION, **W. P. Jones.** *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 88–99.

An Electronic System for Air Traffic Control Information

R. M. KALB
MEMBER AIEE

Synopsis: A digital information processing system has been built for handling and storing flight plans and weather information. Messages are received and sent out by the system automatically over standard teletype lines. Requests for stored information and directions to change it are also transmitted by teletype. This system is capable of being connected in common with other similar systems by a teletype network with automatic switching facilities. It was designed to form part of the common system for air traffic control. Received messages are randomly filed in a magnetic drum store and subsequently found by means of their contents. The system has a complete logical plan which is designed to give a specific result for each possible action and to cause any questionable or obviously erroneous received messages to be printed out for supervisory review.

A NEW kind of electronic system has been developed as part of an over-all, nation-wide air-traffic control plan[1,2] to accommodate the increasingly pressing problem of dispatching air traffic. In recent years, this problem has been aggravated by a growing volume of all sorts of traffic to be monitored—military, private, and commercial—and planes operating at higher speeds have reduced the time available for action. The terminal control problem is complicated by greater landing speeds and by the widening range of speeds of approaching aircraft. At one end of the speed range are found helicopters, at the other end jet airplanes. The trend to all-weather flying imposes a severe terminal problem when planes are landing and taking off under poor weather conditions. The problem of handling today's air traffic, and that of tomorrow, is being resolved under the planning of a common system for air traffic control.[3,4] The common system includes navigational aids, airborne equipment, and ground-based electronic aids for the men who control the air traffic—the controllers. The subject equipment of this paper, a flight plan storage and processing equipment, provides, as part of these ground-based aids, a high-speed, unindexed, electronic filing system.

Control of Air Traffic

Controllers who monitor the cross-country air traffic are located in Air Route Traffic Control (ARTC) centers, about 25 in number. These communicate with airplane pilots and dispatchers through about 500 stations. Three-fourths of these stations handle communications only, the other fourth are control towers at important airports. The geographical area handled by one ARTC center varies greatly according to the density of surrounding air traffic and may contain anywhere from two to seven control towers. The controllers stationed in control towers are concerned only with immediate departures and arrivals of planes at their airport. In handling a flight from take-off to touchdown, both kinds of controllers take action but those at the ARTC centers exercise the continuing monitoring control.

The electronic system makes information on all pertinent flights quickly available to a controller when the time for his action approaches, and otherwise on demand. It receives and delivers weather reports automatically and keeps a file of the flight plans for all air traffic within an area. Table I lists the contents of these types of messages. Routine processing of the information is

---

Paper 55-147, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing December 16, 1954.

R. M. KALB is with Remington Rand, Inc., St. Paul, Minn.

The system described was developed for the Air Navigation Development Board, Civil Aeronautics Administration, Technical Development and Evaluation Center, under contract *C13ca-410*.

Vol. 7
Pt. 1
1955

done automatically. Facilities are provided for altering the filed information and for keeping it up to date. A permanent record is made of all the information handled in the system.

FLIGHT PLANS

Air-traffic information filed is in the form of flight plans. A flight plan is an itinerary of an aircraft plus pertinent data about the craft. Airways, like highways, are designated by route codes. An airway code is a color and a number, e.g., *green 14* or *red 9*. Thus, it is possible to construct an itinerary by a sequence of codes and reporting points, or "fixes," with the altitude of each portion of the route also given. Such a plan is originated by the pilot or by a dispatcher and may be presented via teletype message, voice, radio, or on a written form. The information is reduced to a standard format and put on teletype lines for introduction into the system. A flight plan received in teletype code can enter the electronic system directly; received in any other form it is relayed through an operator who prepares the proper teletype message.

Each flight plan message comprises the following portions:

1. Teletype address signals for automatic line switching and machine clearing (6 characters).
2. A preamble consisting of the address from which the message is sent and codes for the kind and length of message (6 characters).
3. A criterion of message selection consisting of the departure time and airport, and the aircraft identity number (18 characters).
4. General information including arrival time and airport and an alternate arrival airport, plane type, instrument equipment, speed, fuel, and priority, and any restrictions applicable to the flight (35 characters).
5. Route information consisting of airway, altitude, fix and time (1 to 9 sets of 17 characters each).
6. Sign-off, figures shift-carriage return-letters shift, to denote the end of the message (3 characters).

Messages handled by the system are tabulated in Table II. They include both flight plans and intructions concerning the disposition of flight plans. Incoming flight plans to be filed constitute one type of message. Flight plans which modify an already existing one are messages of a second type. Instructions may be to read out a particular flight plan or to cancel one or to activate one whose departure time is imminent; these three kinds of instructions form three more types of message. The system includes provision for storing weather reports as well as flight plans, and the filing, reading out, and cancelling of such reports necessitate three more message types. A reply or an alarm message is generated in each case to complete a cycle of action by the system.

Nature of the Automatic Operations

In processing a received message of any type, the system attempts a sequence of four logical steps. These steps, with their logical alternatives, are diagrammed for the filing process in Fig. 1. The first step is to verify the format of the message in certain respects to make sure that the message is properly composed for the system. The next step is to search the file of stored information. This search is made according to the nature of the incoming message, either to guard against accepting a duplicate message or to find a stored message which is to be modified or cancelled. The third logical step is disposal of the message. This step may be to store the message or to read out, modify, or cancel one already stored. Details of this step are given in Fig. 2. The final step is to prepare a reply giving external indication that the action initiated has been completed or that it cannot be.

Failure of the system to comply or its inability to proceed with any logical step initiates an alarm message. If the format of the incoming message fails verification, the message as received is printed

Table I. Contents of Messages Filed

Characters	Flight Plans		Weather Reports
6	signals for teletype line switching and machine clearing		same
6	preamble:	address	same
		kind of message	
		length of message	
18	criterion:	departure time	time of observation
		departure airport	place of observation
		aircraft identity	the word "weather"
35	general information:	arrival time	text of report
		arrival airport	
		alternate arrival airport	
		type of aircraft	
		speed of aircraft	
		fuel, hours	
		instrument equipment	
		priority of flight	
		restrictions on course	
17 (1 to 9 sets)	route information:	airway	
		altitude	
		fix and time	
3	sign off:	figures shift	same
		carriage return	
		letters shift	

Table II. List of Messages Processed by the System

Nature of Action to Be Initiated	Contents of Incoming Message	Contents of Normal Outgoing Message	Contents of Outgoing Alarm Message*
File flight plan†	complete FP	complete FP	conflicting preamble and criterion
Activate FP	time signal	complete FP	none
Modify FP	complete FP	complete FP	none
Read out FP	criterion and preamble	complete FP	received preamble and criterion
Cancel FP	criterion and preamble	complete FP‡	received preamble and criterion
File weather	complete report	complete report	none
Read out weather	criterion and preamble	complete report	received preamble and criterion
Cancel weather	criterion and preamble	complete report‡	received preamble and criterion

* For logical inconsistencies. If format is erroneous, the received text is relayed to supervisory controller.
† FP stands for flight plan.
‡ To legal storage.

Table III. Data on Storage System

Capacity	2,000 messages	1,250 short, 115 characters maximum 750 long, 230 characters maximum
	330 tracks	285 message store 30 recirculating registers 15 timing and format control
Speed	1,190 rpm drum	22 inches diameter, 30 inches long
	114-kc pulse repetition frequency	for timing circuits
	22.8-kc pulse repetition frequency	for read/write circuits
	410 milliseconds maximum search time	8 drum revolutions
Size	$31^1/_2$ feet long $2^1/_2$ to 3 feet wide 7 feet high	plus 3 by 5 foot console and 2 cooling units each $2^1/_2$ by $3^1/_2$ feet and $7^1/_2$ feet high total floor space 15 by 25 feet
	1,988 tubes	*7AK7* and *5963* types predominate
Power	27.2 kva	includes 2.2 kva for cooling units

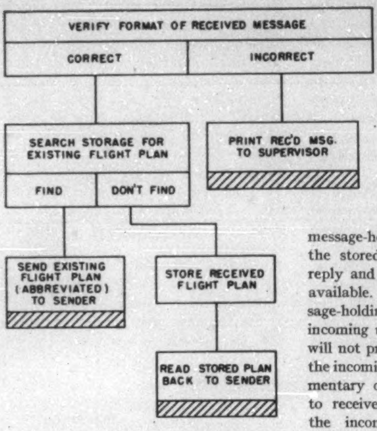

VERIFY FORMAT OF RECEIVED MESSAGE
CORRECT
INCORRECT
SEARCH STORAGE FOR EXISTING FLIGHT PLAN
FIND
DON'T FIND
PRINT REC'D MSG. TO SUPERVISOR
SEND EXISTING FLIGHT PLAN (ABBREVIATED) TO SENDER
STORE RECEIVED FLIGHT PLAN
READ STORED PLAN BACK TO SENDER

Fig. 1 (left). Logical sequence of filing a flight plan

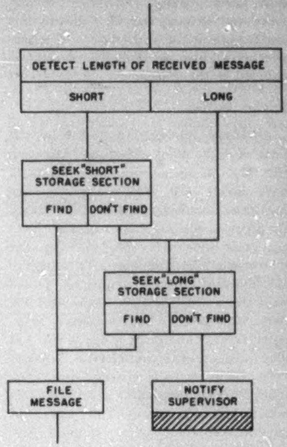

DETECT LENGTH OF RECEIVED MESSAGE
SHORT
LONG
SEEK "SHORT" STORAGE SECTION
FIND
DON'T FIND
SEEK "LONG" STORAGE SECTION
FIND
DON'T FIND
FILE MESSAGE
NOTIFY SUPERVISOR

Fig. 2 (right). Subsidiary logic of storing operation

out to a supervisor at the machine center who then handles the situation by the exercise of his own judgment. All other alarms result in the composition of a message containing only a preamble and criterion. If a flight plan received for filing has the same aircraft identity number and departure airport as a plan already stored, that plan is rejected and the preamble and criterion of the plan in file are transmitted to the sender who thereby is advised of the conflicting plan; he then can take further action by cancelling this plan before filing the new one or by having all of the filed plan read out to him as the basis for deciding further action.

As soon as a message is filed initially, the complete text of the message is read back from storage and transmitted to the sender. This gives him at once a verification and a typed copy of the flight plan he has on file. When any message in file is cancelled, it thereupon is read out in its entirety to a tape perforator; the record of this perforator, termed legal storage, constitutes a complete listing of all messages upon which action has been completed. Tape from the perforator can later be fed into a page printer to reproduce a typed copy of any desired portion of the record. The system deletes spaces from the received message before storing the information contained in it, and restores the spaces as the message is read out. This action is based on a fixed format for each type of message.

Buffer registers, called incoming message-holding units, receive messages as they arrive and introduce them into the storage system when it is ready after they have been verified. An outgoing message-holding unit similarly accepts the stored information necessary for a reply and relays it to a teletype line as available. Incoming and outgoing message-holding units operate in pairs; an incoming unit will accept a message but will not process it into the system until the incoming unit has acquired a complementary outgoing message-holding unit to receive the reply. The preamble of the incoming message is transferred directly to the outgoing message-holding unit for utilization by it in formulating the reply message.

A high-speed message-holding unit is provided also. This is a dual unit which handles information both ways between the storage system and a display panel console at speeds one order of magnitude greater than in the teletype lines.

A special feature of the system is a clock-actuated read out. By action of this device, the storage is searched completely every 10 minutes and all plans for flights due to depart within the succeeding 20 minutes are selected automatically and transmitted to the airport of departure. When a flight plan is so activated, a status-indexing character filed with it by the system is changed to indicate "enroute status." This indication prevents the same flight plan from being read out automatically again during a subsequent periodic review. Provision is made also for filing plans with an enroute status initially. This is done when the plan is filed within 10 minutes of its departure time or when the controller modifies a plan already under his scrutiny.

Properties of the Equipment

Pertinent numerical characteristics of the equipment are given in Table III.

Courtesy Civil Aeronautics Administration, Technical Development and Evaluation Center

Fig. 3. Flight plan storage and processing equipment

The filing system utilizes magnetic drum storage with a capacity of 2,000 flight plans and weather reports. These are grouped into two categories, short and long. A short message consists of not more than 115 characters, a long message can be twice as long. Information is stored in the teletype code with the characters serial and their five bits in parallel. The total information storage capacity of the drum is 316,250 characters or 1,581,250 bits.

Stored information is so arranged that during the search only the criterion portion of each message is scanned. With this arrangement, the entire contents of the information store can be searched in eight drum evolutions, approximately two-fifths of a second.

The timing pulse rate of the drum is 114 kc. Using a fivefold interlace, information is read and written at about 23 kc. Conversion between the teletype line speed of 460 characters per minute and internal system speeds of 23,000 characters per second is accomplished through the incoming and outgoing message holding units. More details of the equipment are published elsewhere.[5]

A complete set of the equipment which forms the system described has been installed for evaluation at the Technical Development and Evaluation Center of the Civil Aeronautics Administration at Indianapolis, Indiana. An over-all view of the installation is shown in Fig. 3. This installation is being connected to operate through an automatic teletype switching center,[6] a modified *81D1* system, and in co-operation with data transfer and display equipment described in a companion paper.[7]

Future Plans

Planning for the air-traffic control system calls for dividing the country into control zones, perhaps 10 or 15 altogether, and centralizing the control for each zone in one similar equipment located within it.[8] When this is done, equipments for adjacent zones would be connected by teletype lines with automatic switching so that two equipments could "converse" with each other to establish flight plans for interzone traffic. The system logic for such operation was examined prior to planning the present system, and it was designed according to principles allowing such interconnection of equipments.

Requirements for the ultimate system will not be known until after the present one is evaluated. One matter to be settled concerns the relative demands for weather and traffic information and whether such information should be filed on one equipment, as in the present instance, or on separate equipments.

Another question to be determined is the possible need for adding a time-reservation facility to the system.[9] The purpose of this facility would be to limit the number of landings and take-offs scheduled during periods of 1 hour or less for each airport. The limit would, of course, have to be variable because of the influence of weather and other conditions which affect the acceptance rate of an airport. The means for including time reservations in the air-traffic control system were worked out during the early planning of the present equipment, and it was so planned as to permit the addition of the time reservation feature to subsequent models.

The system described constitutes an initial venture into automatic teletype filing and processing of alphabetical and numerical information. It is also the first serious attempt to mechanize the routine of handling air-traffic control information. With the possibility realized of machines composing messages and conversing via teletype lines, facilities are at hand to further the advance of clerical procedures by means of electronics.

References

1. Air Traffic Control, Special Committee 31. *Paper 27-48/DO-12*, Radio Technical Commission for Aeronautics, Washington, D. C., May 12, 1948.

2. System of Air Navigation and Traffic Control Recommended by the Radio Technical Commission for Aeronautics, P. C. Sandretto. *Electrical Communication*, New York, N. Y., vol. 26, March 1949, pp. 17–27.

3. Air Traffic Control and the National Security, Air Coordinating Committee. *U. S. Government Printing Office*, Washington, D. C., Dec. 1950, pp. 51–65, 69–77.

4. The Air Traffic Story, Special Committee 47. *Paper 194-52/DO-47*, Radio Technical Commission for Aeronautics, Washington, D. C., Dec. 1, 1952.

5. Message Storage and Processing with a Magnetic Drum System, A. P. Hendrickson, G. I. Williams, J. L. Hill. *Proceedings*, Eastern Joint Computer Conference, Philadelphia, Pa., Dec. 1954

6. A Fully Automatic Private-Line Teletypewriter Switching System, W. M. Bacon, G. A. Locke. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 473–80.

7. Data Transfer and Display Equipment for a Proposed System of Air Traffic Control, G. E. Fenimore. *AIEE Transactions*, vol. 74, pt. I, May 1955, pp. 145–49.

8. Airport Time Utilization Equipment, Phase I, A Systems Engineering Study, D. E. Olshevsky. *Report No. JA-627-P-1*, Cornell Aeronautical Laboratory, Inc., Buffalo, N. Y., March 1950, pp. 20–32, 83–87.

9. *Ibid*, pp. 1–19, 50–60.

Machine Testing for Deviation of Data from a Poisson Distribution

F. A. HADDEN
ASSOCIATE MEMBER AIEE

THERE are many situations in which data are obtained describing events that occur at random. If the mechanism causing the random events is such that the events are independent and the probability of occurrence in any given interval of time is constant, the occurrence is described by the Poisson distribution. In a population of events which are normally expected to occur according to the Poisson distribution, certain events may occur as the result of special factors which operate in relation to them only. In such cases, these specially influenced events will not fall within the Poisson distribution that describes the rest of the population. It may be desirable to study the factors causing their peculiar occurrence and this will require that the normally occurring events be separated from the normally occurring ones. A practical machine method for making this kind of separation for any data which are Poisson-distributed is outlined. For clarity, the techniques will be described in terms of a particular example but other applications are of course possible.

A type of event which is usually Poisson-distributed is the in-flight failure of vacuum tubes in a piece of air-borne electronic equipment. The data in this case could be the number of failures of each tube type in the same time period, as

---

Paper 55-154, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing December 9, 1954.

F. A. Hadden is with The Rand Corporation, Santa Monica, Calif.

Table I. Vacuum-Tube Failure Data

Tube Type	Failures	Sockets
2D21	53	5
2J56	159	1
2K25	8	1
4C35	22	1
5R4	47	4
5Y3	11	1
6AC7	9	1
6AK5	118	10
6AL5	228	19
6AU6	40	4
6C4	10	1
6J6	310	25
6Y6	145	6
12AU7	76	7
12AX7	293	24

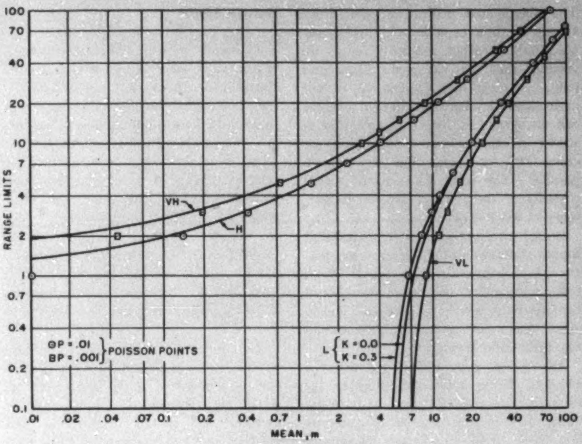


Fig. 1. Poisson approximation curves

Table II. Data and Calculations

Tube Type	Number of Failures	Sockets	Expected Mean	H	L	VH	Adjusted Expected Mean	Adjusted H	Adjusted L	Test Result
2D21	53	5	69.5	90.0	50.4		59.5	78.6	41.8	
2J56	159	1	13.9	23.7*	5.5	27.0*				VH
2K25	8	1	13.9	23.7	5.5		11.9	21.0	4.2	
4C35	22	1	13.9	23.7	5.5		11.9	21.0*	4.2	H
5R4	47	4	55.6	74.1	38.5		47.6	64.8	31.8	
5Y3	11	1	13.9	23.7	5.5		11.9	21.0	4.2	
6AC7	9	1	13.9	23.7	5.5		11.9	21.0	4.2	
6AK5	118	10	139	168	112		119	145	93.9	
6AL5	228	19	264	302	227		226	262	191	
6AU6	40	4	55.6	74.1	38.5		47.6	64.8	31.8	
6C4	10	1	13.9	23.7	5.5		11.9	21.0	4.2	
6J6	310	25	348	393	305		298	339	258	
6Y6	145	6	83.4	106*	62.4	134*				VH
12AU7	76	7	97.3	121	74.6		83.3	106	62.3	
12AX7	293	24	334	378	292		286	326	247	
	1529	110								

First Mean Rate $=\frac{1{,}529}{110}=13.9.$

Adjusted Mean Rate $=\frac{1{,}529-159-145}{110-1-6}=\frac{1{,}225}{103}=11.9.$

* Tested out of range at this point.

shown in Table I. These data will be used as the example throughout the paper. The purpose of the testing described is to check the circuit application and tube design criteria for each tube type to make sure that the tubes are neither overloaded nor underloaded relative to the other types in the equipment. The problem is to identify the tube types having failure probabilities significantly different from those for the average tube type.

Since failures occur at random, it is expected that there will be a variation in the number of failures of each tube type. The average type will occasionally have a number of failures which is either unusually high or unusually low, while the abnormal type will generally have an unusual number of failures because of design or other differences. The separation test must be able to distinguish between the average and abnormal types without too many mistakes. The test described is based on determining from the mean of the data at hand the range of numbers of failures that will include the average tube types a reasonable fraction of the time.

For the Poisson distribution, there is no simple equation relating the mean number of failures m of the average tube types and the upper limit of the range H and lower limit of the range L; but the relation can be plotted as a graph or can be approximated by simple equations. From Poisson tables, such as those by Molina,[1] the total (cumulative) probability of there being H or more failures for any given mean m can be obtained as well as the total (cumulative) probability of there being L or fewer failures for any given mean m. For use in testing, the approximate relations between H and m and L and m for various fixed values of probability are needed. Experience has shown that for identifying a category, e.g., a tube type not belonging to the common Poisson distribution it is reasonable to use two probability values such as 1/100 and 1/1000. These two probabilities define the four curves: very low (VL), low (L), high (H), and very high (VH). All four curves are a function of m alone.

For many applications of the technique presented here, such as the work on equipment failures[2] which prompted this paper, a category which has only one occurrence, however improbable, does not warrant investigation except in very special cases which can be treated separately. The curves and equations given here have been adjusted so that two is the lowest number of occurrences which will test as greater than H.

Since the Poisson and Gaussian distributions approach each other for large values of the mean, a Gaussian distribution with the same mean, standard deviation, and cumulative probabilities as the Poisson is an appropriate basis for approximation. A good fit was obtained to each of the desired curves relating the range limits to the mean for use with the machine procedures outlined in the following. The four equations which are satisfactory, for all positive values of m are

$$VL=m-3.09\sqrt{m}+1.2$$

$$L=m-2.33\sqrt{m}+K$$

$$H=m+2.33\sqrt{m}+1.1$$

$$VH=m+3.09\sqrt{m}+1.6$$

In Fig. 1, the actual Poisson points for the probabilities 1/100 and 1/1,000 are the dots in circles and squares respectively. The approximation equations given in the foregoing are the plotted curves; there are two curves for L, one with $K=0.3$ and one with $K=0$. The $K=0.3$ approximation is a very close fit to the 0.01 (or 1/100) probability all the way, while the $K=0$ approximation deviates a little for low m and L; e.g., the probability is about 0.007 for $L=3$ and 0.004 for $L=0$. The choice of a value for K is discussed in the following. The constraint of never accepting an H less than two is convenient for this approximation in the middle range of m, but it does raise H slightly for the lower range of m. The square root required in the equations may be avoided by using the ratio of two polynomials in m as an approximation, but this is limited to a restricted range of values for m.

These equations give the limits within which the number of failures of an average tube type will almost always fall, based on the expected mean number of failures for the tube type. If all tube types appear in the same number of sockets, then the expected mean and the range limits will be the same for all the tube types; but if the number of sockets varies with the tube type, then the expected means will be proportional to the number of sockets, with the range limits varying correspondingly.

In the example given in Table II, there are 15 tube types in 110 sockets with a total of 1,529 failures. The failure rate per socket is

$$\frac{1{,}529}{110}=13.9$$

or the expected mean is the per-socket rate times the number of sockets that the tube appears in, e.g., $13.9\times5=69.5$ for the *2D21* tube and $13.9\times1=13.9$ for the *2J56* tube, etc.

For each tube type, the range limits are found from the equations or the curves. It is seen that the *6J6*, which has the largest number of failures, is actually within limits whereas the *2J56* with a moderate number of failures is above the *VH*-range limit.

Caution must always be used in the testing of categories which have had unequal exposure to the mechanism causing the occurrences. In the example, the unequal exposure was caused by the variation in the number of sockets using each type of tube. In another kind of analysis, it might be caused by one piece of equipment operating more hours than another. In general, however, the basic rate from which the expected mean for each category is determined is the number of occurrences per unit of equipment per unit of exposure.

If the data include failures from abnormal tube types as well as from the average type, the value of the mean calculated from all the data can be shifted by the failures of the abnormal types. One extremely high failing type can mask a few moderately high failing types and even force a few average types which happen by chance to be a little low to appear abnormally low. In the example, the *4C35* was masked by the *2J56* and the *6Y6* until the data were adjusted. The amount of care required to protect against these effects will depend upon the character of the data. The detailed procedures described here are for the simple case with few abnormal types and small penalty for minor errors. More complicated data can be handled by a relatively simple extension of the adjustment techniques.

Machine Procedure Outline

For the procedure to be outlined it is assumed that each category, e.g., tube type, to be compared is represented by one accounting machine card such as those used in International Business Machines. These cards are referred to as category cards. Each category card has punched in it the number of occurrences, e.g., failures, and the exposure, e.g., number of sockets, for that category in appropriate units. In addition to the category cards there are two summary cards. One of these, called the first summary card, has the total number of occurrences for all categories and the total exposure for all categories punched in it. The first summary card is placed in front of the category cards and the blank second summary card is placed behind. The cards are passed in this order through a calculating punch:

From the first summary card, store and retain for further use the total number of occurrences for all categories and the total exposure for all categories. Calculate the mean rate of occurrence per category per unit exposure and store this quantity for further use. As each category card passes, multiply the exposure for that category by the mean rate of occurrence to obtain the expected mean number of occurrences for that category. Using the equations given, calculate L and H, test the number of occurrences for that category to see if the number falls outside the 1/100 range and then, if the category does not test outside, pass to the next category card. In calculating L, the constant $K=0.3$ should be used unless insufficient capacity is available in the calculating punch to make the number of occurrences test as outside the range on the L end for m less than 1.

If the category does test outside, calculate the appropriate *VH* or *VL* and test to see if the number of occurrences falls outside the 1/1000 range. In any case, subtract the number of occurrences for that category from the total number of occurrences stored and subtract the exposure for that category from the total exposure stored. The purpose of these subtractions is to eliminate those categories that are known to differ from the average distribution so that in subsequent iterations they will not affect the mean rate which describes the average categories. Punch the card to describe the way it falls outside the ranges and pass to the next category card.

Punch the adjusted total number of occurrences and total exposure into the second summary card which follows all of the category cards. Other groups of category and summary cards can follow and be tested automatically without any change in the machine. After all groups being tested have made the first pass, they are ready for the first iteration. Replace the first summary card by the second summary card and replace the old second summary card by a new second summary card. Repeat the preceding steps but ignore the already marked cards.

The procedure is repeated as many times as necessary but some experience has shown that even one iteration may be sufficient for many applications. More complicated data can be handled by eliminating only *VH* and *VL* categories on the first pass or by eliminating only the one most extreme *VH* or *VL* category each pass until all the *VH* and *VL* categories are gone. It is not necessary to rearrange and add summary cards physically as just described because the same effect can be achieved by properly locating extra blank cards to be used as the new second summary cards when the old ones are advanced to first summary cards. The replacement of a first by a second summary card only requires that the cards be turned over and passed through the machine in the reverse order, using a different machine control board which takes this into account.

References

1. Poisson's Exponential Binomial Limit (book), **E. C. Molina.** D. Van Nostrand Book Company, Inc., New York, N. Y., 1952.

2. Electronic Reliability and Supply Improvement Based on Failure Reporting and Presentation, **F. A. Hadden, L. W. Sepmeyer.** *Rand Corporation Research Memorandum RM-1257*, The Rand Corporation, Santa Monica, Calif., May 20, 1954.

A Tapered Strip Transmission Line for Pulse Transformer Service

F. G. PRIMOZICH
ASSOCIATE MEMBER AIEE

E. R. SCHATZ
ASSOCIATE MEMBER AIEE

J. B. WOODFORD
ASSOCIATE MEMBER AIEE

THE need for wide-band impedance-transforming devices has long been recognized. The increasing interest in the use of very short pulses of millimicrosecond duration in radar, communication, and other systems accentuates the demand for a device which is capable of faithfully transforming such pulses. As the bandwidth and upper cutoff frequency increase, it becomes more and more difficult to construct suitable lumped-parameter or iron-core transformers of conventional design. Thus, the trend in recent years has been toward distributed systems, particularly of the tapered transmission line variety. Various tapers such as the linear,[1] Bessel,[2,3] exponential,[4-8] and hyperbolic[9] have been proposed and analyzed by several investigators. In addition, many different physical structures employing these tapers have been constructed and tested experimentally. These were the 2-wire open line,[4,5] the concentric circular cylinders,[7-9] and the shielded wire and tape helices.[10,11] By and large, each experimental model has been physically bulky and difficult to construct. It is the purpose of this paper to investigate a tapered transmission line design which is reduced in physical bulk and which can be fabricated more economically.

The Rectangular Strip Transmission Line

A structure which can readily be applied to the design of a tapered transmission line is the rectangular strip line. A rectangular strip transmission line consists of a thin conducting strip of width $2a$ which is suspended coaxially by insulating material of permeability μ_1 and permittivity ϵ_1 in a conducting right rectangular cylinder of width $2c$ and height $2b$, which serves as the outside sheath. The cross section of this line is shown in Fig. 1.

The rectangular strip line can be either completely enclosed, as shown in Fig. 1, or the perpendicular sides can be eliminated, as shown in Fig. 2. These two lines are essentially identical in their electrical characteristics over a wide range of ratios of the dimensions. In this treatment, the rectangular strip line without sides, Fig. 2, will be used exclusively. However, all conclusions can also be applied directly to the line with sides.

It has been shown by other authors[10] that a tapered transmission line can be designed by assuming that the characteristic impedance at a point can be determined from the equations derived for a uniform line. This is a valid assumption, provided the taper is not very rapid.

As shown in the Appendix, the phase velocity of a uniform rectangular strip line is equal to $1/\sqrt{\mu_1\epsilon_1}$ and the characteristic impedance is

$$Z_0=\sqrt{\frac{\mu_1}{\epsilon_1}}\,\frac{1}{4}\,\frac{K}{K'} \tag{1}$$

where K and K' are complete elliptic integrals of modulus k. Equations for finding k are given in the Appendix.

By use of these equations and the tables,[12,13] families of curves of normalized characteristic impedance $Z_0\sqrt{k_r/\mu_r}$ versus the dimensional ratio a/c are plotted for various values of the ratio b/c. The resulting curves are shown in Fig. 3. Here $k_r=\epsilon_1/\epsilon_0$ and $\mu_r=\mu_1/\mu_0$, where ϵ_0 and μ_0 are the values for free space.

With an air dielectric the range of impedance levels readily obtainable is 20 to 300 ohms. It should be noted, however, that any magnitude of impedance is theoretically possible since the curves approach the ordinate asymtotically. The only practical limitation is the ability to control accurately the dimensions of the inner strip as the ratio a/c approaches zero. The impedance curves can be raised or lowered by proper choice of ϵ_1 and μ_1.

Design of an Exponential Strip Line

The steady-state and transient characteristics of the exponential transmission line have been widely discussed.[4,5,7,8,10] Schatz and Williams[8] have determined the pulse response of the exponential line and have established the criteria for the design of such lines as pulse transformers, i.e., length and flare requirements for the transformation of a particular pulse of known shape and duration have been given. This study will concern itself mainly with the design of a rectangular strip line with exponential taper although, with the curves given in Fig. 3, any form of taper could be considered.

In an exponential line the characteristic impedance at a point varies as an exponential function of the distance from the sending end

$$Z_x=Z_0\epsilon^{\gamma x} \tag{2}$$

where

γ = flare coefficient which is assumed positive
x = distance from the sending end of line
Z_0 = characteristic impedance at the sending end of line

After the length and flare requirements have been established for a particular pulse application, it is only necessary to taper the center conductor (the strip) of a rectangular strip line in such a way that the characteristic impedance along the line will follow that given in equation 2. From equation 2 the impedance at a given position x on the line is known. For given values of b and c, the value of a necessary to give the required value of

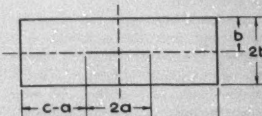


Fig. 1. Cross section of completely enclosed strip line

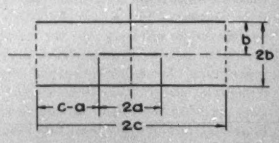


Fig. 2. Cross section of strip line without sides

Vol.
Pt. 1
195

Paper 55-177, recommended by the AIEE Special Communications Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted June 14, 1954; made available for printing December 3, 1954.

F. G. PRIMOZICH is with the North American Aviation, Inc., Downey, Calif., and E. R. SCHATZ and J. B. WOODFORD are with the Carnegie Institute of Technology, Pittsburgh, Pa.

This paper contains part of a dissertation by F. G. Primozich in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Carnegie Institute of Technology. The authors wish to acknowledge the support of Office of Naval Research Contract *N7onr30308*. F. G. Primozich also wishes to acknowledge the support of the Buhl Foundation through a Buhl Fellowship.

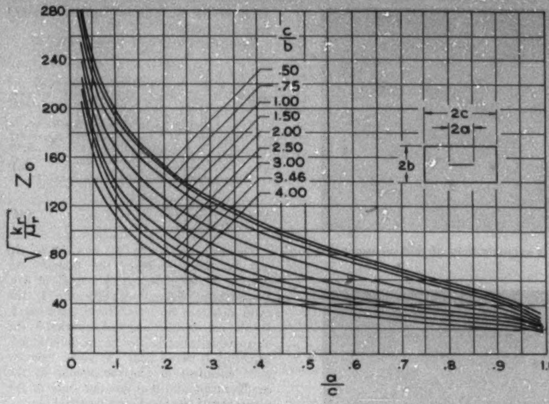


Fig. 3. Normalized characteristic impedance of the rectangular strip line in ohms

impedance at the point x can be determined from Fig. 3; i.e., a as a function of x, distance along the line, is now known and the inner conductor is then tapered accordingly. The dimensions at the sending end of the line are, of course, functions of desired impedance level, of space limitations, and of breakdown and power-handling requirements.

One of the particular space-saving possibilities of the tapered rectangular strip line is that it can be conveniently rolled up; see Fig. 4. This particular model was designed to be 10 meters long with an impedance transformation of 4. The impedance level at the low end is 25 ohms and at the high end 100 ohms. The total dimensions of the line shown in Fig. 4 are about 2 by 4 feet. For convenience, in rolling, this line was constructed without sides, but it would be a simple matter to add sides in any commercial fabrication.

Pulse Response of the Strip Line

An experimental determination of the pulse response of the line in Fig. 4 was made. The line was terminated with an impedance equal to the impedance level at the output. The input pulse is shown in Fig. 5 and the output pulse in Fig. 6. The length of pulse at the input is about 36 millimicroseconds. The output pulse is to the same scale.

Conclusions

It can be stated that the rectangular strip line is particularly suited to tapered transmission line design. Because of the flat shape it can be conveniently rolled or placed against a flat supporting surface; therefore, such a line is definitely a space saving device. In addition, because of the simple construction, tapering of the inner conductor is easy. Printed circuit techniques might also be used in the manufacture of such a line, greatly reducing the complexity of fabrication. All in all, this type of line is a space-saving, cost-reducing device which can help make the use of tapered transmission lines more practical.

Appendix. Phase Velocity and Characteristic Impedance

If the cross section shown in Fig. 1 is transformed conformally into a section of an infinite parallel plate transmission line, the characteristic impedance and phase velocity can be determined. The capacitance and inductance per unit length as well as the characteristic impedance and phase velocity of a parallel plate line are well known.[13] A section of the parallel plate line is shown in Fig. 7.

$$c/l=\epsilon_1\frac{d}{e}=\text{capacitance per unit length} \qquad (3)$$

$$L/l=\mu_1\frac{e}{d}=\text{inductance per unit length} \qquad (4)$$

$$Z_0=\sqrt{\frac{\mu_1}{\epsilon_1}}\,\frac{e}{d}=\text{characteristic impedance} \qquad (5)$$

$$v=\frac{1}{\sqrt{\mu_1\epsilon_1}}=\text{phase velocity} \qquad (6)$$

Since the capacitance and inductance are invariant in a conformal transformation, the characteristic impedance and phase velocity determined for the transformed line also apply to the original transmission line. Now the required transformation must be performed. Since the cross section shown in Fig. 1 is symmetrical, the problem may be solved in one quarter of the section; see Fig. 8.

This section can be transformed into one of four identical parallel plate lines that are electrically in parallel. If e is the separation and d is the width of an equivalent parallel strip line, the section in Fig. 8 will transform into the section shown in Fig. 9. The object of the transformation is to obtain the dimensions e and d in terms of a, b and c.

It has been shown[14] that from a Schwartz-Cristoffel transformation which transforms Fig. 8 in the Z plane into Fig. 9 in the W plane the following relationships can be derived

$$\frac{K_0}{K_0'}=\frac{c}{b} \qquad (7)$$

$$\frac{K}{K'}=\frac{4e}{d} \qquad (8)$$

The subscript 0 denotes the Z plane. K and K' are complete elliptic integrals of modulus k, which is given by the following relationship

$$k^2=\frac{[sn^2(k_0,Z_4)-sn^2(k_0,Z_3)][sn^2(k_0,Z_1)-sn^2(k_0,Z_2)]}{[sn^2(k_0,Z_1)-sn^2(k_0,Z_3)][sn^2(k_0,Z_4)-sn^2(k_0,Z_2)]} \qquad (9)$$

where Z_1, Z_2, Z_3, and Z_4 are the co-ordinates of the points, as indicated in Fig. 8.

$sn(k_0, Z)$ is a Jacobian elliptic function with modulus k_0 to which the following definitions apply.[15] Let

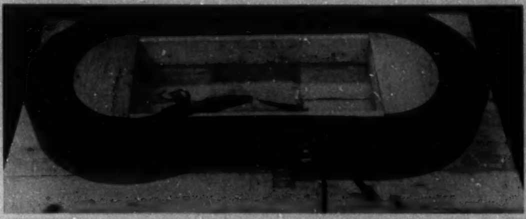


Fig. 4. Rolled-up tapered strip line pulse transformer

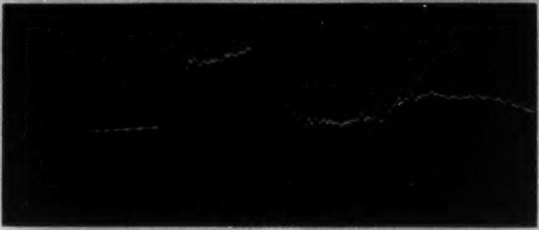


Fig. 5. Input to pulse transformer. Pulse width is approximately 36 millimicroseconds

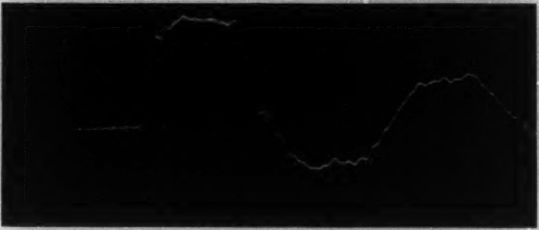


Fig. 6. Output of pulse transformer

$$\mu=\int_0^{\varphi}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}(k^2<1) \tag{10}$$

$$=\int_0^{q}\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}(t=\sin\varphi) \tag{11}$$

where

$\mu=F(\varphi, k)$ = elliptic integral of the first kind
φ = amplitude
k = modulus

$$\sin\varphi=sn(k,\mu) \tag{12}$$

$$\cos\varphi=cn(k,\mu)=\sqrt{1-sn^2(k,\mu)} \tag{13}$$

$$dn(k,\mu)=\sqrt{1-k^2sn^2(k,\mu)} \tag{14}$$

The complementary modulus is

$$k'=\sqrt{1-k^2} \tag{15}$$

$$K=\int_0^{\pi/2}\frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}=\text{complete elliptic integral} \tag{16}$$

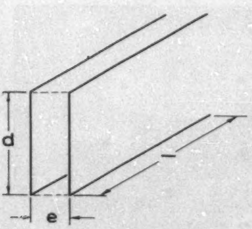


Fig. 7. Parallel plate line

$$K'=\int_0^{\pi/2}\frac{d\varphi}{\sqrt{1-(k')^2\sin^2\varphi}}=\text{complete elliptic integral of the complementary modulus} \tag{17}$$

A complete discussion of the properties of Jacobian elliptic functions appears in references 16 and 17.

The following are some relations which can be used in the evaluation of the functions[18]

$sn(k, 0)=0$ $\qquad sn(k, K)=1$
$cn(k, 0)=1$ $\qquad cn(k, K)=0$
$dn(k, 0)=1$ $\qquad dn(k, K)=k'$

$$\begin{aligned} sn[k, (K+jK')]&=1/k \\ cn[k, (K+jK')]&=jk'/k \\ dn[k, (K+jK')]&=0 \end{aligned} \tag{18}$$

In Fig. 8 it can be seen that if $Z_1=0$, $Z_2=K_0(1-a/c)$, $Z_3=K_0$, and $Z_4=K_0+jK_0'$, then

$$sn(k_0, Z_1)=sn(k_0, 0)=0 \tag{19}$$

$$sn(k_0, Z_3)=sn(k_0, K_0)=1 \tag{20}$$

$$sn(k_0, Z_4)=sn(k_0, K_0+jK_0')=1/k_0 \tag{21}$$

If the ratio $c/b=K_0/K_0'$ is known, a value of the modulus k_0^2 can be obtained from the tables.[12]

Equation 9 can now be simplified to

$$k^2=\frac{\left[\left(\frac{1}{k_0}\right)^2-1\right]\left[-sn^2(k_0, Z_2)\right]}{\left[-1\right]\left[\left(\frac{1}{k_0}\right)^2-sn^2(k_0, Z_2)\right]} \tag{22}$$

$$k^2=\frac{sn^2(k_0, Z_2)(1-k_0^2)}{1-k_0^2sn^2(k_0, Z_2)} \tag{23}$$

$$k^2=\frac{(k_0')^2sn(k_0, Z_2)}{dn(k_0, Z_2)} \tag{24}$$

The values of $sn(k_0, Z_2)$ and $dn(k_0, Z_2)$ can be obtained from the tables[12] for various values of k_0 or c/b and Z_2 or a/c. Once these values are determined, the ratio $4e/d$ can be found from equation 8. By the use of this ratio and equation 5, the characteristic impedance can be determined. Equation 6 shows that the phase velocity is independent of the dimensions of the configuration and is a function only of the constants of the insulating material.

It is of interest to note that the characteristic impedance of a strip line as shown in Fig. 2, but in which the outer conductors extend to infinity, is essentially identical to the characteristic impedance of the enclosed line of Fig. 1 for all values of c/b which are plotted when a/b is less than 0.4 and is also for the entire range of a/b for values of c/b greater than 3. Thus, it might be concluded that the completely enclosed line is essentially identical in this range with the line with the finite-plate outer conductor.

References

1. SINUSOIDAL CURRENTS IN LINEARLY TAPERED LOADED TRANSMISSION LINES, **John W. Arnold, Paul F. Bechberger.** *Proceedings*, Institute of Radio Engineers, New York. N. Y., vol. 19, Feb. 1931, pp. 304–10.

2. NON-UNIFORM LUMPED ELECTRIC LINES, **S. Ballantine.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 203, April 1927, pp. 561–82.

3. THE NON-UNIFORM TRANSMISSION LINE, **A. T. Starr.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 20, June, 1932, pp. 1,052–63.

4. THE EXPONENTIAL TRANSMISSION LINE, **C. R. Burrows.** *Bell System Technical Journal*, vol. 17, Oct. 1938, pp. 555–73.

5. TRANSMISSION LINES WITH EXPONENTIAL TAPER, **H. A. Wheeler.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 27, Jan. 1939, pp. 65–71.

6. THE TAPERED TRANSMISSION LINE, **J. W. Milnor.** *AIEE Transactions* (*Electrical Engineering*), vol. 64, June, 1945, pp. 345–46.

7. PULSE TRANSIENTS IN EXPONENTIAL TRANSMISSION LINES, **E. R. Schatz, E. M. Williams.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 38, Oct. 1950, pp. 1208–12.

8. DESIGN OF EXPONENTIAL LINE PULSE TRANS-

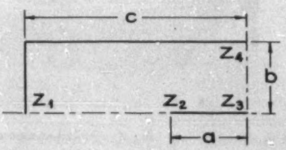


Fig. 8 (left). Quarter-section strip line

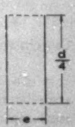


Fig. 9 (right). Parallel plate line

formers, E. M. Williams, E. R. Schatz. *Ibid.*, vol. 39, Jan. 1951, pp. 84–86.

9. The Hyperbolic Transmission Line as a Matching Section, Herbert J. Scott. *Ibid.*, vol. 41, Nov. 1953, pp. 1654–57.

10. Exponential-Line Section Pulse Transformers, Joseph Kukel. *Ph.D. Thesis*, Carnegie Institute of Technology, Pittsburgh, Pa., 1953.

11. A Broadband Transition From Coaxial Line to Helix, C. O. Lunt. *RCA Review*, Radio Corporation of America, Princeton, N. J., vol. 11, March 1950, pp. 133–42.

12. Tafeln der Besselshen, Theta-, Kugel- und Anderer Funktionen (book), Keiichi Hayashi. Julius Springer, Berlin, Germany, 1930.

13. Smithsonian Elliptic Functions Tables (book), G. W. Spenceley, R. M. Spenceley. *Smithsonian Institution*, Washington, D. C., vol. 109, 1947.

14. Current Flow in Rectangular Conductors, H. Fletcher Moulton. *Proceedings*, London Mathematical Society, London, England, vol. III, Jan. 1905, pp. 104–10.

15. Tables of Integrals and Other Mathematical Data (book), H. B. Dwight. The Macmillan Company, New York, N. Y., 1947.

16. Jacobian Elliptic Functions (book), E. H. Neville. Oxford University Press, London, England, 1951.

17. The Applications of Elliptic Functions (book), A. G. Greenhill. The Macmillan Company, New York, N. Y., 1892.

18. An Introduction to the Theory of Functions of a Complex Variable (book), E. T. Copson. Oxford University Press, London, England, 1950, p. 384.

19. Reference Data for Radio Engineers (book). Federal Telephone and Radio Corporation, 3rd ed., 1949, p. 325.

Cold-Cathode Counting Circuits

H. L. FOOTE
ASSOCIATE MEMBER AIEE

IN RECENT years, considerable emphasis has been placed on high-speed counting devices for use in electronic computers and designed to operate at pulse rates of the order of 100 kc, and higher. These of necessity involved hot-cathode electron tubes.

There are, however, certain counting and computing problems to be met which do not require the extreme speeds of the complex computers. Several low-speed (up to 700 pulses per second (pps)) counting devices have been evolved using cold-cathode gas triodes. These tubes have the advantage of requiring no heater power with attendant thermal shock, and of having a life dependent upon the number of on-off operations rather than upon hours of operation. The *RCA-5823* tubes used in all of the counters to be described have a minimum life of approximately 40,000,000 operations. Tubes chosen at random and placed on life test in the Stromberg-Carlson laboratory have not failed even after 250,000,000 operations. The standard heater-type tube has a life of about 2,000 hours, with premium types ranging up to 10,000 hours.

Cold-Cathode Gas Tubes

Although this discussion will be limited to 3-element cold-cathode tubes, it should be noted that 2-element and multielement types are available for special purposes. The three elements, Fig. 1, are the main anode *A*, the cathode *K*, and the starting anode *SA*. A sufficient potential difference between the starting anode and the cathode causes the gas in the tube to ionize and the anode-to-cathode gap to conduct. A tube remains ionized until the potential from anode to cathode is reduced below the sustaining value.

In the *5823*, the starting anode must be raised to a minimum of +75 volts above the cathode potential to fire the tube, first ionizing the starting-anode-to-cathode gap, then transferring the discharge to the main-anode-to-cathode gap. The main-anode potential must be above the sustaining value of the main gap of about +85 volts, but not above the self-ionizing value of about +290 volts. Once the main-anode-to-cathode gap is ionized, the starting anode has no further control over conduction. The tube is extinguished or de-ionized by opening either the main-anode or the cathode circuit.

The speed of response of the tube is dependent upon the ionization time (the time necessary to transfer the starting-gap discharge to the main gap) and the de-ionization time. The former is of the order of 20 microseconds, the latter about 500 microseconds.

In general, the main anode is connected to a positive voltage source, and the cathode is connected to ground through a load resistance. A positive potential (usually a short pulse), sufficient to ionize the starting anode-to-cathode gap is applied to the starting anode. A d-c bias potential somewhat lower than this required voltage may be applied to the starting anode, thus reducing the amplitude of the ionizing pulse required.

Decade Counting

The first of two decimal counting chains to be described is shown in Fig. 2. In this circuit,[1] to limit current drain, each tube in firing extinguishes the preceding tube in the counting chain. This is accomplished by using a common main-anode current-limiting resistance *R*. At the instant a tube fires, current through it is limited only by *R*, with the result that the potential of all anodes drops to about +85 volts. Since the cathode of the preceding tube is held momentarily to about +75 volts by its cathode capacitor, there is no longer sufficient voltage to maintain the main gap ionization, and the tube is extinguished.

Start tube *ST* provides the first tube in the counter *V*1 with an enabling bias derived from the divider in the cathode circuit, until the first pulse to be counted arrives. The start tube is ionized by the application of the anode voltage to the circuit. Its cathode capacitor charges, raising the starting anode bias on *V*1 to a value sufficient to allow the first pulse to fire *V*1. When *V*1 fires, the start tube is extinguished by the process described earlier.

Pulse rates and pulse widths are important in choosing the cathode capacitor values, since the operation of the counter depends on the charge time of this capacitor. The bias potential derived from the preceding tube must not be great enough to fire a tube with no pulse present, and conversely, the pulses must not be of sufficient amplitude to fire a tube without the enabling bias. Chains of

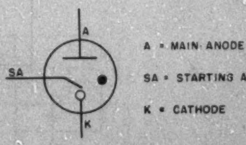


Fig. 1. Cold-cathode triode

Paper 55-197, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted March 26, 1954; made available for printing November 24, 1954.

H. L. Foote is with the Stromberg-Carlson Company, Rochester, N. Y.

The circuits of Figs. 1, 3, and 5 to 7 were the work of Stromberg-Carlson Company engineers R. B. Trousdale, G. R. Paul, and the author. Figs. 2 and 4 are described in reference 1.

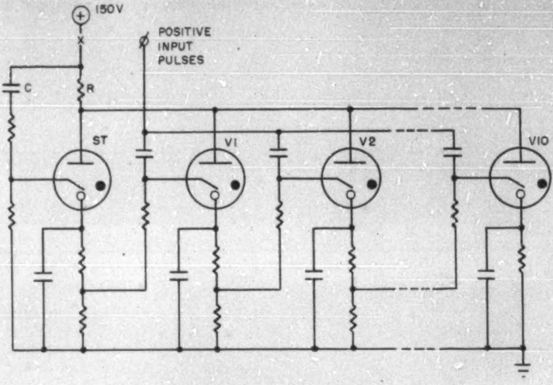


Fig. 2. Cold-cathode decimal counting chain

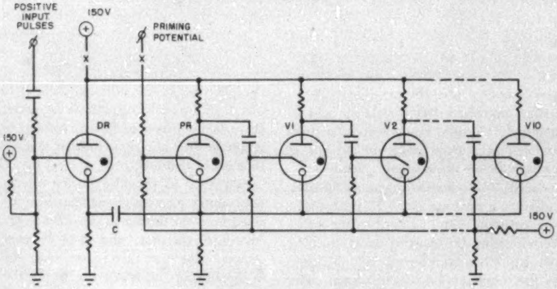


Fig. 3. Cold-cathode decimal counting chain

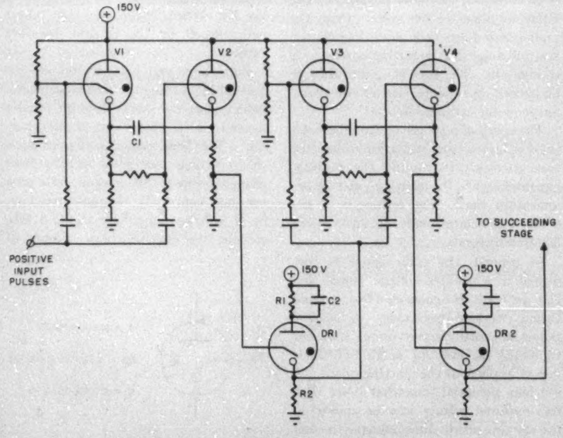


Fig. 4. Cold-cathode binary counting chain

this type have been designed to count at rates up to 200 pps.

A second decade counter, Fig. 3, using a fixed starting-anode bias, employs a common drive tube *DR* so that the pulse source is not loaded by parallel inputs to the counting tubes as in the previously described chain. In this chain also, a start, or prime, tube *PR* is ionized prior to counting. The individual anode and the common cathode resistors are of equal value. Therefore, with a 65-volt drop in the tube, the cathode bus will rise to +42.5 volts. The starting anodes of all of the tubes in the chain are biased to +90 volts. Since the common cathode potential is at +42.5, the starting anode-to-cathode potential is still not great enough to ionize the tube. The starting anode of the drive tube is biased to +50 volts. When a pulse of sufficient magnitude appears on this starting anode, the tube will be ionized raising the cathode potential momentarily +85 volts. At this time, three events occur. The common cathode bus rises to +85 volts as capacitor *C* charges. Since the anode potential of *PR* is +107.5 volts, the tube can no longer maintain ionization and will go out. The drive tube, meanwhile, extinguishes itself by relaxation because of the large cathode resistance. When *PR* goes out, its anode rises sharply from +107.5 volts to +150 volts. This positive pulse appears on the starting anode of *V*1 and adds sufficiently to the fixed bias to fire *V*1. This process is repeated as additional positive pulses are applied to the starting anode of the drive tube.

Pulse repetition rates and pulse widths again determine the value of capacitor *C*. The drive tube must be extinguished with the accompanying common cathode bus potential drop before the succeeding tube can fire. There is, therefore, a time race between the extinguishment of *V*1 and the drive tube and the appearance of the positive pulse on the starting anode of *V*2. Chains of this type have been made to count at pulse rates up to 100 pps.

Binary Counting

For many applications, counting in binary form is preferable to decimal counting because of the savings in tubes that are possible where large numbers are involved. In binary notation a number is described using only two symbols, 0 and 1. Since 2 to the zero power=1, it may readily be seen that any number can be expressed as the sum of powers of two, just as any number in the decimal system can be expressed as the sum of powers of ten. The first ten integers given in the binary scale follow.

In the first binary chain to be considered, Fig. 4,[1] the starting anode of tube *DR*1 receives sufficient voltage to ionize the tube when capacitor *C*1 coupling the cathodes of *V*1 and *V*2 is charged, and *V*1 fires.

Decimal	Binary
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
10	1010

This occurs on all odd input pulses except the first. Initially, the full supply voltage is applied to the main anode of *DR*1 but, as capacitor *C*2 becomes charged, this voltage decreases. The resistance of *R*1 +*R*2 is great enough to cause *DR*1 to deionize as soon as the starting anode potential falls below the critical point. A sharp pulse appears at the cathode of *DR*1 each time *V*1 fires, and the input to the second counting stage consists of positive pulses at a repetition rate half that of the input to the first stage.

The functioning of the individual stages is as follows: With *V*1 ionized, the starting anode of *V*2 is biased to +50 volts. A positive pulse on the starting anode will then fire *V*2. Capacitor *C*1 is charged to +85 volts because of the ionization of *V*1: as *V*2 fires, the potential at its cathode rises to +85 volts, increasing the potential at the cathode of *V*1 to +170 volts, and *V*1 is extinguished. When *V*1 is ionized again, *V*2 will be extinguished in a similar manner.

This chain is difficult to read directly, and a modification has been designed using two tubes per binary digit, for 0 and 1 representation. In addition, the 0 tubes in each stage must be primed on prior to the start of counting. The chain has been used at pulse rates up to 100 pps.

A simpler type of binary counting chain has been developed which will count at pulse rates up to nearly 700 pps. This circuit is shown in Fig. 5. Here again two tubes per stage are used to indicate 0 and 1, with the 0 tubes primed on prior to counting. The two tubes in stage one are supplied with a fixed bias by divider *R*1-*R*2 so that they can be driven by the input pulses. When *V*2 fires, its cathode potential provides the bias for the second stage. The second pulse ionizes *V*1 and causes *V*2 to be extinguished by the process described for the

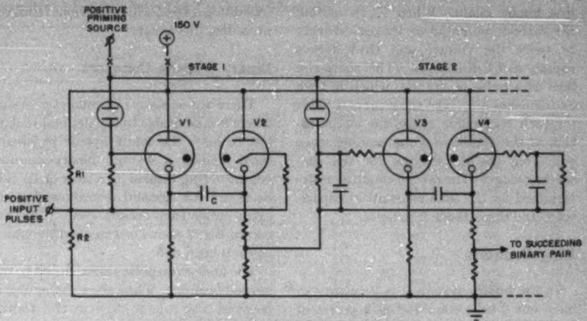

POSITIVE PRIMING SOURCE
150 V
STAGE 1
STAGE 2
V1
V2
V3
V4
R1
R2
C
POSITIVE INPUT PULSES
TO SUCCEEDING BINARY PAIR

Fig. 5. Cold-cathode binary counting chain

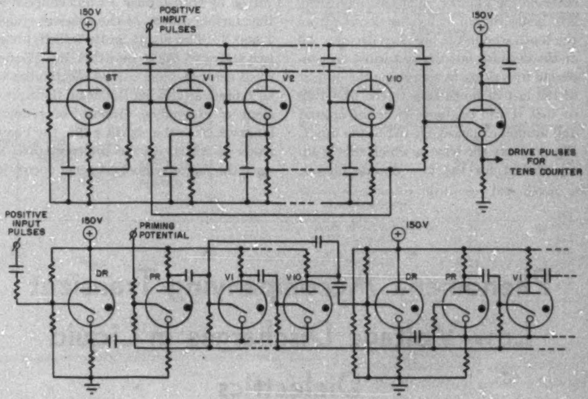

150V
POSITIVE INPUT PULSES
ST
V1
V2
V10
150V
DRIVE PULSES FOR TENS COUNTER
POSITIVE INPUT PULSES
150V
PRIMING POTENTIAL
DR
PR
V1
V10
150V
DR
PR
V1

Fig. 6. Decade ring circuits

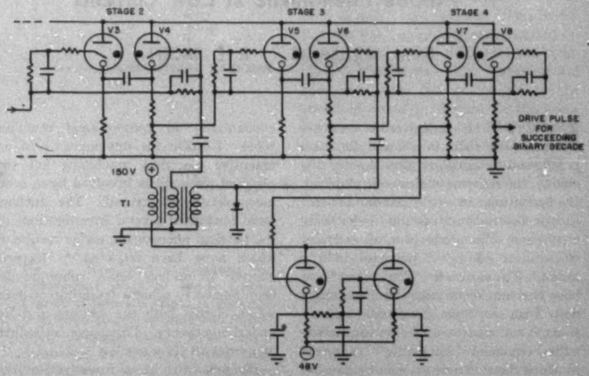

STAGE 2
STAGE 3
STAGE 4
V3
V4
V5
V6
V7
V8
DRIVE PULSE FOR SUCCEEDING BINARY DECADE
150 V
T1
48V

Fig. 7. Binary decade counter

first binary chain. When $V1$ is ionized, the cathode potential of $V2$ rises sharply to twice the voltage and then decays rapidly as $C1$ discharges. This sharp rise acts as a pulse to fire $V4$, which in turn extinguishes $V3$. The same process holds for each succeeding stage, the de-ionization of the 1 tube in the preceding stage providing the drive pulse for the succeeding stage. The rate of counting is determined by the time constant of capacitor C and the cathode resistance.

Ring Counters

In the binary system it is necessary to have one 2-tube stage for each power of two required to describe the number. In decimal counting, a decade counter is required for each power of ten required to describe the number, each succeeding decade being driven by a pulse derived from the tenth step in the previous decade. All of the decades but the last must be converted into rings, by returning the output of the last or tenth step to the first step so that it will continue to count around and around as pulses are fed to the input. Fig. 6 shows the way in which this is accomplished for the two decade chains. Circuit operation will be obvious from the preceding discussion.

Binary Decade Counters

There are occasions in computer design where it is desirable to use decimal scaling but to have the count appear in binary form. Since a 4-stage binary counter counts in ring fashion on a scale of 16, it is necessary to account somehow for six extra pulses each decade cycle. The circuit of Fig. 7 shows one method by which this has been done.

The first seven pulses are counted in the normal fashion. When the eighth pulse is received and the 0 tube in the fourth stage is extinguished, its sharp cathode pulse is used to trigger a 2-stage delay circuit, the output of which is phase inverted by transformer $T1$ and coupled to the starting anodes of the tubes in stages 2 and 3. The eighth pulse returns these two stages to the 0 condition, but the derived delayed pulse causes the 1 tubes to be refired, advancing the chain the equivalent of six pulses. Instead of reading binary 8 after the eighth pulse, the chain reads 14 which can be interpreted as 8, since the single symbol 14 never occurs in the scale of 10. Similarly, binary 15 is interpreted as 9, and the binary zero as decimal zero. The stored count is as follows:

Decimal	Binary
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000, shift to 1110
9	1111
0	0000

The cathode pulse obtained when the 1-tube of stage four is extinguished in going from binary 15 to zero, is used to drive the succeeding binary decade.

While the paper is limited to counting devices, it will be realized that cold-cathode tubes are adaptable to other low-speed computing circuits such as multipliers, pulse generators, delay networks, and memories.

Reference

1. THE DESIGN OF SWITCHING CIRCUITS (book), William Keister, Alistair E. Ritchie, Seth H. Washburn. D. Van Nostrand Company, New York, N. Y., 1951.

Phenomena Accompanying Transient Low-Voltage Discharges in Liquid Dielectrics

I—Anode Phenomena at Low Currents

E. M. WILLIAMS
MEMBER AIEE

R. E. SMITH
ASSOCIATE MEMBER AIEE

ELECTRICAL spark-over between metallic electrodes in a liquid dielectric is followed by a high current-density discharge, the removal of electrode material, the formation of well-defined and characteristic electrode craters and, under some conditions, the production of colloidal suspensions of metal particles in the liquid. The associated phenomena have been the subject of many studies, originally with emphasis on colloidal suspensions,[1,2] and more recently, in connection with "electrical" machining[3,4] processes in which the electrode material removal is adapted for purposes analogous to those encountered in conventional machine tools. Despite the development of an extensive electrical machining art the physical phenomena involved have been incompletely understood. The authors have conducted several investigations of the physical phenomena, earlier results of which have been reported.[3,4] Experimental techniques have subsequently been refined to achieve quantitative data which clarify both the process and its underlying theory. This paper is the first of a series on these results.

The first effort in the investigation was directed to the simplification of experimental conditions. In the electric spark-over[4] machining process a series of high-current low-voltage discharges between a tool (cathode) electrode and a workpiece (anode) electrode gradually wear the workpiece into a desired form. Although all the significant phenomena should be observed in the wear caused by a single discharge, most experimental studies have been concerned with measurements of the cumulative effect of a very large number of discharges over a sustained period.

When the effects of many discharges in rapid sequence are superposed, numerous factors combine to obscure the physical phenomena; among these are known[4] to be redeposition of eroded material, both cathode and anode, formation of low-pressure vapor pockets, and temperature effects due either to the dissipation of large amounts of average power or

---

Paper 55-72, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing December 8, 1954.

E. M. WILLIAMS and R. E. SMITH are with the Carnegie Institute of Technology, Pittsburgh, Pa.

The work described in this paper was supported in part by the Method X Company, affiliate of the Firth Sterling Corporation, Pittsburgh, Pa.

Table I. Properties of Test Specimens Used as Anode Electrodes

Specimen	Alloy Designation	Resistivity ×10⁻⁶ Ohm-Centimeter	Density, P/Inches³	Tensile Strength K Psi	Specific Heat Cal/gm/C*	Heat of Fusion Cal/gm	Heat of Vaporization Cal/gm	Melting Point C	Boiling Point C	Brinell Hardness
Tungsten	Pure	5.56	0.690	597	0.032	61	960	3,370	6,000	290
Molybdenum	Pure	5.70	0.325	270	0.061	70	1,460	2,620	4,700	250
18-8-SMO steel	316	76.7	0.29	125	0.12	69.8	1,528	1,385	2,830	160
Hot rolled monel		54.0	0.318	96	0.13	68	1,313	1,330	2,830	162
FC steel	416	60.5	0.28	91	0.15	68.6	1,551	1,495	2,820	165
18-8-SM steel	303	76.8	0.29	81	0.12	70.0	1,530	1,425	2,820	140
17ST aluminum	17ST	5.56	0.102	64	0.23	91.5	2,513	645	2,490	110
Free cutting brass	240	7.95	0.307	61	0.09	39.6	931	885	1,910	131
Hard phosphor bronze B		10.10	0.321	60	0.09	46.7	1,210	1,050	2,490	128
Naval brass	386	7.23	0.304	75	0.09	39.8	928	900	1,860	136
Soft copper	204	1.76	0.323	32	0.092	48.9	1,250	1,087	2,500	42
Magnesium	Pure	4.86	0.063	28	0.25	85.7	1,315	651	1,103	35
Zinc	Pure	6.66	0.258	24	0.09	26.3	440	420	907	33
2S aluminum	2S	3.12	0.098	19	0.23	93.5	2,580	660	2,500	42
Cadmium	Pure	7.59	0.312	12	0.055	13.6	212	320	768	33
Lead	Pure	21.9	0.398	2.2	0.03	5.75	206	327	1,750	3.8
Tungsten carbide	HAX3		0.570	150				2,800		

* Calories per gram per C.

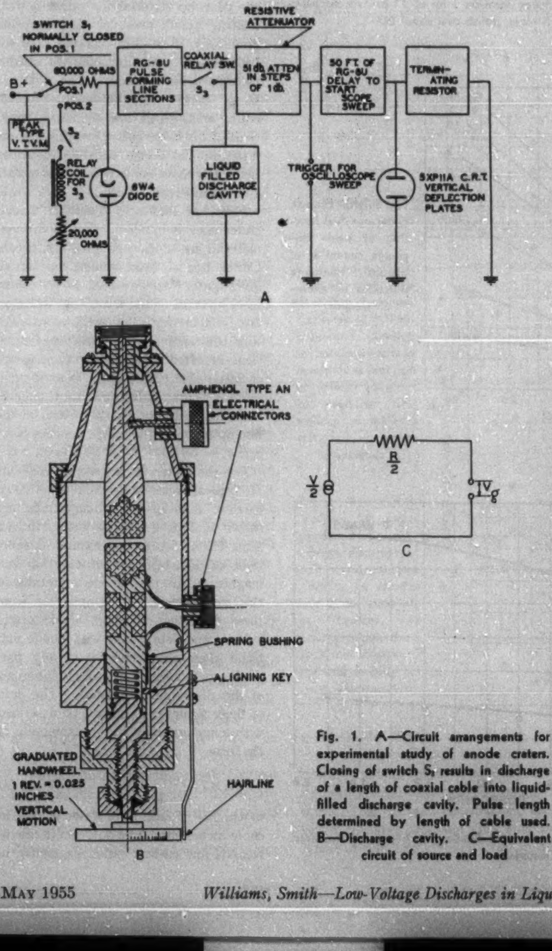

RESISTIVE ATTENUATOR
SWITCH S_1 NORMALLY CLOSED IN POS. 1
B+
80,000 OHMS
POS. 1
POS. 2
PEAK TYPE V.T.V.M.
S_2
RELAY COIL FOR S_3
6W4 DIODE
20,000 OHMS
RG-8U PULSE FORMING LINE SECTIONS
COAXIAL RELAY SW.
S_3
51 db ATTEN IN STEPS OF 1 db
LIQUID FILLED DISCHARGE CAVITY
50 FT. OF RG-8U DELAY TO START SCOPE SWEEP
TRIGGER FOR OSCILLOSCOPE SWEEP
TERMINATING RESISTOR
5XP11A C.R.T. VERTICAL DEFLECTION PLATES
A
AMPHENOL TYPE AN ELECTRICAL CONNECTORS
SPRING BUSHING
ALIGNING KEY
GRADUATED HANDWHEEL 1 REV. = 0.025 INCHES VERTICAL MOTION
HAIRLINE
B
R/2
V/2
V_g
C

Fig. 1. A—Circuit arrangements for experimental study of anode craters. Closing of switch S_1 results in discharge of a length of coaxial cable into liquid-filled discharge cavity. Pulse length determined by length of cable used. B—Discharge cavity. C—Equivalent circuit of source and load

Table II. Diameters of Anode Craters Produced by 11-Ampere 1.5-Microsecond Discharge

Anode Material	Crater Area ×10⁻⁴ Inch²	Melting Point, C
Group 1—Diameter Dependent on Anode Material		
Lead	6.24	327
Cadmium	5.15	320
Zinc	3.01	420
Magnesium	2.74	651
Group 2—Diameter Relatively Independent of Anode Material		
Soft copper	2.32	1,087
Hard phosphor bronze	2.30	1,050
17ST aluminum	2.27	645
Naval brass	2.17	900
18–8 SM steel	2.17	1,425
Free cutting brass	2.27	885
316 steel	2.29	1,385

the formation of continuous discharges.[4] Furthermore, although emphasis has been placed on the distinction between anode and cathode phenomena, the discharges used in experimental studies have been oscillatory so that each electrode serves both as anode and cathode. This investigation was accordingly concentrated on the effects of single, unidirectional time-limited discharges. The experimental conditions were further limited to the effects of discharge currents in the lower range of those encountered in electrical machining processes, for reasons which will subsequently become apparent.

The test procedure involved the initiation of spark-over between prepared anode and cathode electrodes at appropriate spacings by application of voltages of the order of 50 to 1,000 volts. The circuit was arranged to deliver, following spark-over, a single current pulse of substantially constant amplitude and controlled duration at convenient values in the range of 17 to 1,500 millimicroseconds. The circuit arrangements are shown in Fig. 1. The liquid dielectric employed

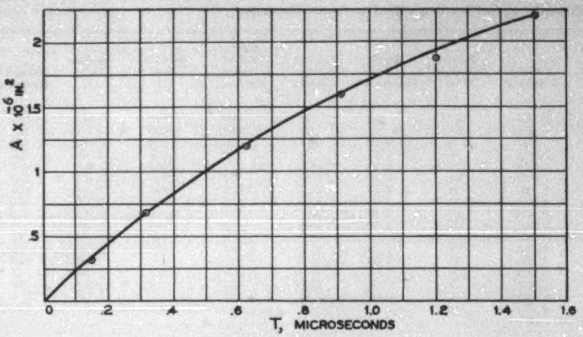


Fig. 2. Anode crater area A as function of discharge duration T for an 11-ampere discharge in materials of group 2 of Table II. Melting points over about 600 C

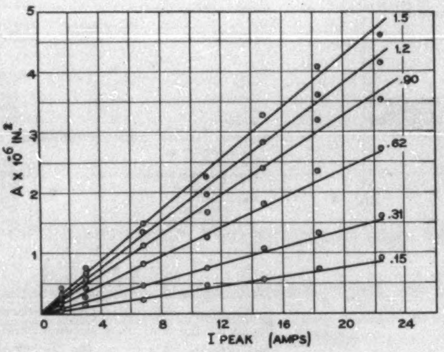


Fig. 3 (left). Anode crater area A as function of peak discharge current I for various discharge durations in microseconds

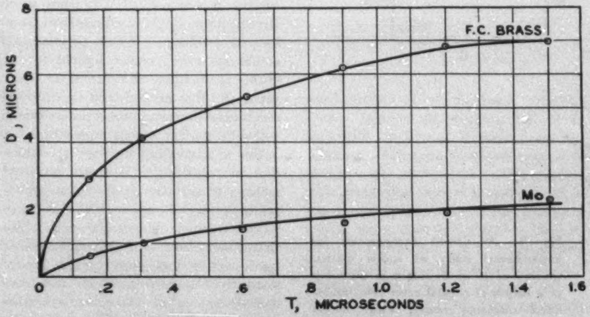


Fig. 4. Anode crater depth D as function of discharge current duration T in 11-ampere discharge for two anode materials

was kerosene from a single bulk sample. Although the chemical composition of commercial kerosene is not exactly defined, it is known that chemical composition of the dielectric has no effect on the nature of the physical phenomena in the discharge. The quantitative effect of the dielectric composition will be discussed in a subsequent paper.

The anode electrodes used were annealed specimens of metals or alloys with the properties shown in Table I. Tensile strength and resistivity values were determined by test while other properties listed are from standard references.

Although a variety of cathode materials was used, it was found that anode crater size and form are independent of the chemical and physical properties of the cathode material for cathodes with melting points over 400 degrees centigrade (C). The data reported here will include tests conducted only with higher melting point cathodes. The data obtained consist of tabulations of anode crater area, depth, and volume as functions of discharge current magnitude and duration and anode electrode constitution. For convenience many of these data are presented graphically.

First, as shown in Table II, anode crater area is independent of the anode material for a wide range of materials. Crater area has been assumed by various investigators[5] to be equal to the cross-sectional area of the electrical discharge; this conclusion is probably warranted with the relatively short discharge durations reported here. The independence of crater area with respect to anode material is thus an indication that the anode is not a significant source of ions in the discharge unless the anode melting point is low.

For the materials of group 2 of Table II, typical crater areas as a function of current duration and magnitude are shown in Figs. 2 and 3. Figs. 4 and 5 show crater depth and volume as functions of discharge current duration and magnitude for two anode materials in this same group. It is apparent from these data that although anode crater volume and depth vary with anode material these characteristics in any particular material are fairly simply functions of the discharge parameters. The data of these figures can, in fact, be described with some accuracy by an expression of the form

$$V = K_1 \times T(I)^{3/2} \qquad (1)$$

in which V, T, and I are crater volume in cubic inches, discharge duration in microseconds and current pulse magnitude in amperes respectively, and K_1, a constant characteristic of the material. Values of K_1 for a number of the materials tested are as follows:

Tungsten, 7.37×10^{-13}.
Molybdenum, 16.29×10^{-13}.
18–8 *SMO* steel, 35.2×10^{-13}.
Hot-rolled monel, 45.8×10^{-13}.
Naval brass, 58.7×10^{-13}.
17ST aluminum, 68.7×10^{-13}.
F. C. brass, 72×10^{-13}.

The relation between the physical properties of the anode material and the crater volume for a particular current magnitude and duration may be expected to cast further light on the controversial question as to the mechanism of crater formation, numerous theories for which have been advanced.[3]

For the purpose of clarifying the significant phenomena, anode crater volumes under typical conditions, for all the materials of Table I, have been plotted in Figs. 6, 7, 8 and 9 against the physical properties that have been assumed most significant by various investigators. Thus in Fig. 6 crater volume is plotted against energy required to vaporize a unit volume of electrode material initially at room temperature. Fig. 7 is a similar display of crater volume with respect to energy required to melt a unit volume initially at room temperature. In Figs. 8 and 9 volume is plotted against the inverse of the melting points and tensile strengths respectively. For the higher melting point materials there is no more than a trend towards smaller volumes as the melting temperature is increased. This trend could be a reflection of the generally higher tensile strengths of higher melting-point materials. Actually a much better correlation exists between crater volume and tensile strength and this correlation is emphasized in Fig. 10, in which crater depth is plotted against the reciprocal of tensile strength. From this curve it may be deduced that the value of K_1 in equation 1 can be computed for materials with melting points over 600 C as, approximately

$$K_1=\frac{0.44\times10^{-8}}{K_T} \qquad (2)$$

in which K_T is the ultimate tensile strength of the material in pounds per square inch. The constant is applicable to discharges in a kerosene dielectric, and, as will appear from data to follow is limited to nonferromagnetic materials.

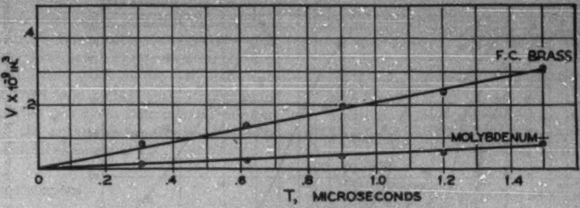


Fig. 5. Anode crater volume V as function of discharge current duration T in 11-ampere discharge for two anode materials

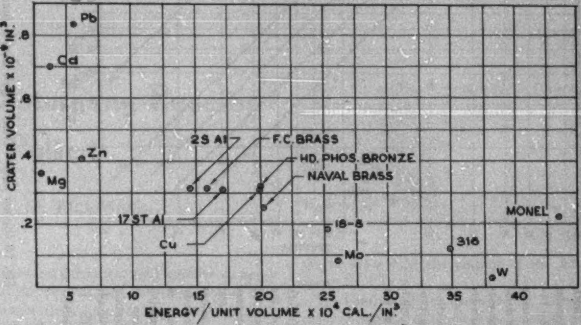


Fig. 6 (above). Volume of crater produced by 11-ampere discharge of 1.5-microsecond duration in anodes of various materials as function of energy required to vaporize unit volume of material initially at room temperature

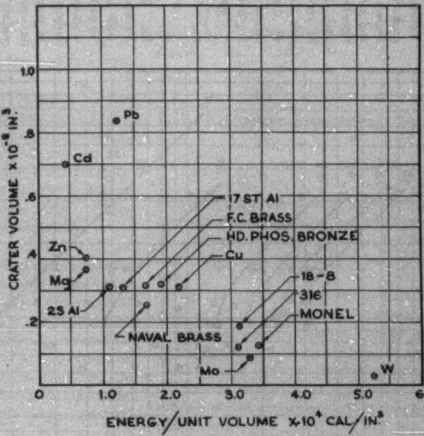


Fig. 7 (right). Volume of crater produced by 11-ampere discharge of 1.5 - microsecond duration in anodes of various materials as function of energy required to melt unit volume of material initially at room temperature

These data lead one to set aside thermal action as the primary factor in anode erosion in materials with melting points above 600 C. A question remains concerning the possibility of explosive force or a cathode "flare" action. It can, however, be demonstrated that the anode crater form is primarily dependent on the current distribution under the anode surface. This current distribution, for a given discharge area, can vary only as a result of skin effect. Skin effect in nonferromagnetic materials should be a negligible factor in the range of fractional microsecond durations of relatively small discharge currents chosen for the tests reported here; this choice was made, in part, for the purpose of minimizing skin effect, therefore emphasizing the role of

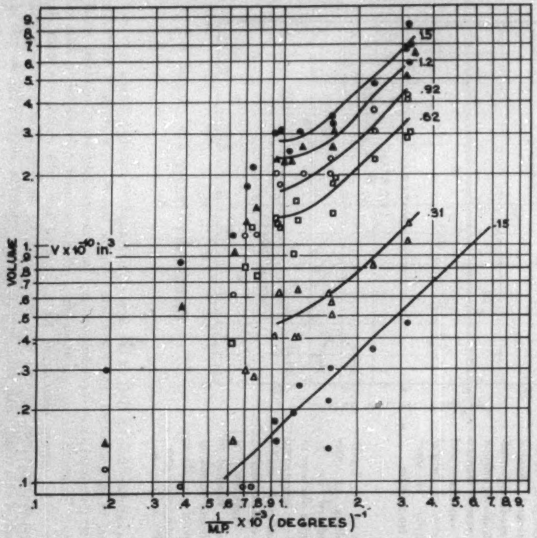
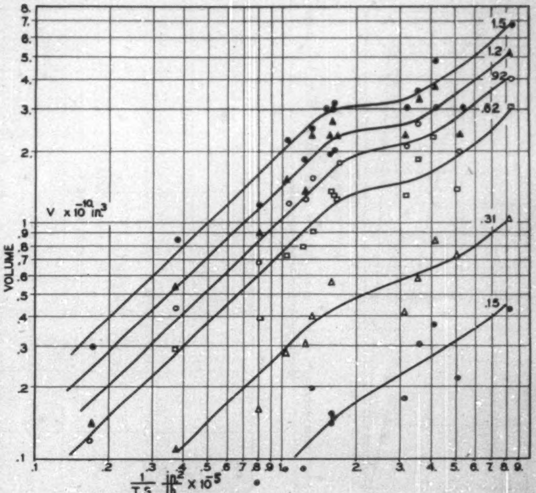


Fig. 8 (left). Volume of crater produced by 11-ampere discharge and various durations in microseconds, in anodes of various materials as a function of the reciprocal of the melting point of the material

Fig. 9 (right). Volume of the crater produced by an 11-ampere discharge and various durations in microseconds, in anodes of various materials as a function of the reciprocal of the tensile strength TS of the material



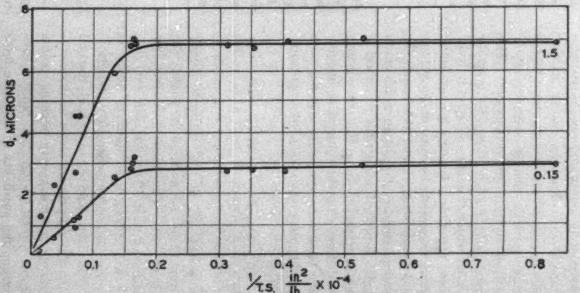
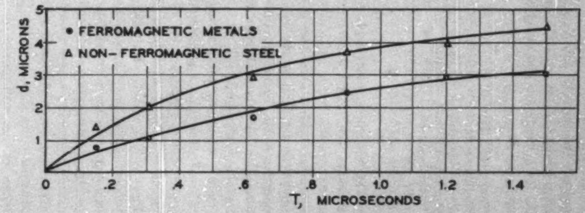


Fig. 10 (below left). Crater depth d as a function of the reciprocal of the anode ultimate tensile strength, at room temperature, for the anode crater produced by H ampere discharges of 0.15- and 1.5-microsecond duration

Fig. 11 (below). Crater depth d as a function of the discharge duration in a discharge of 11 amperes for ferromagnetic and nonferromagnetic materials of approximately equal tensile strength



physical properties other than conductivity. However, with a ferromagnetic material, skin effect is appreciable and the current magnitude below the anode surface is substantially decreased. The effect of a changed current distribution on the anode crater is illustrated in Fig. 11 for ferromagnetic and nonferromagnetic anodes of approximately equal tensile strength.

These data supply some confirmation for an electric field force hypothesis.[3] Moreover, quantitative data have been advanced previously to show that an ionic force can be sufficient to cause rupture.[3,4] It is possible that explosive forces play some part. Electrochemical phenomena would seem to be excluded since current (and consequently charge transported in a given time interval) appears in the expression for the crater volume with an exponent of 1.5.

References

1. Darstellung Kolloidaler Metallösungen durch Elektrische Zerstäubung, G. Bredig. *Zeit Angew Chemie*, vol. I, 1898, pp. 951–54.

2. Die Natur der Elektrischen Kolloidsynthesis, Carl Benedicks. *Arkiv fur Mathematik Astronomi och Fysikuigefuet of K. Svenska Vetenskapsakademien.* Stockholm, Sweden, vol. 8, no. 7, 1912.

3. Theory of Electric Spark Machining, Everard M. Williams. *AIEE Transactions*, vol. 71, pt. II, Jan. 1952, pp. 105–08.

4. Recent Developments in the Theory and Design of Electric Spark Machine Tools, E. M. Williams, J. B. Woodford, Jr., Richard E. Smith. *AIEE Transactions*, vol. 73, pt. II, May 1954, pp. 83–88.

5. Current Densities in the Cathode Spots of Transient Arcs, S. M. Somerville, W. R. Blevin. *Physical Review*, New York, N. Y., vol. 76, no. 9, 1949, p. 982.

6. Formation of Metallic Bridges between Separated Contacts, G. L. Pearson. *Ibid.*, vol. 56, 1939, p. 471.

Economic Design of Saturating Reactor Magnetic Pulsers

R. A. MATHIAS
ASSOCIATE MEMBER AIEE

E. M. WILLIAMS
MEMBER AIEE

l. 74
1
55

SATURABLE reactors have been widely applied in the electric control art. A less well-known use is in pulse-shaping or pulse-switching circuits, examples of which are to be found in ignitron trigger circuits and in modulators for high-power oscillators in radar transmitters. In the latter instance, circuits may become quite complex and there has been a need for criteria which will result in a simple design procedure. In this paper some useful relations applicable to this problem are described.

The circuit of a typical saturating reactor, or pulsactor, pulser for radar,[1] often termed a Melville line, is shown in Fig. 1. This circuit operates in the following general manner. While the source voltage rises initially, the capacitor C_1 charges through the reactor L_0. The constants are chosen so that the voltage on C_1 rises to a high value during the first cycle. For most of this voltage swing L_1 is unsaturated so that its voltage drop is approximately equal the voltage across C_1. The design of L_1 is such that it reaches a saturated state when the voltage on C_1 reaches a peak. C_1 then rapidly discharges its energy into C_2 at a rate determined by the saturated inductance of L_1. C_2 is chosen so that its peak charge is reached at the instant the energies in C_1 and L_1 are zero. C_2 is equal to C_1 in an ideal case in which there are no core losses and the following stages do not appreciably load the $C_1L_1C_2$ loop. At the instant C_2 reaches peak, voltage L_2 saturates and permits a rapid transfer of energy to the next capacitor. This process continues from loop to loop. Although all capacitors are equal, successive reactors are designed for successively reduced saturated inductance so that, by the time the energy has reached the last stage, it can be discharged into the load in a very short time interval.

The line thus acts to compress the energy transfer interval so that peak power transferred increases from stage to stage. In the ideal case (negligible loading and losses), the peak voltage from stage to stage remains constant and the peak current increases with each stage. Quite useful information on general design problems can, in fact, be obtained from an analysis based on the ideal case especially for designs using the rectangular loop materials now available. This analysis will now be made, with the further assumption that there are no interstage transformers.

The time to charge the kth capacitor C_k will be equated to the time for the kth pulsactor to go through its maximum flux swing to a saturated value. Accordingly, for centimeter-gram-second units

$$t_{k-1}=\pi\sqrt{L_{sk-1}\frac{C}{2}}=\frac{2\Delta B}{E_m}N_kA_k10^{-8} \tag{1}$$

but

$$L_{sk}=\frac{N_k^2A_k4\pi\mu_s10^{-9}}{l_k}=\frac{(N_kA_k)^24\pi\mu_s10^{-9}}{M_k} \tag{2}$$

where $M_k=A_kl_k$, the volume of the kth core.

By combining equations 1 and 2, the ratio of saturated inductances is found to be

$$\frac{L_{sk}}{L_{sk-1}}=\frac{\left(\frac{C}{2}E_m^2\right)}{M_k}\frac{\pi^2\mu_s10^7}{(\Delta B)^2}=\frac{W_m}{M_k}G \tag{3}$$

where

$W_m=\frac{C}{2}E_m^2$, the energy per pulse

$G=\frac{\pi^2\mu_s10^7}{(\Delta B)^2}$, a constant for the reactor

Equation 3 for the ratio of saturated inductances is fundamental to the design of magnetic pulsers of this type and, furthermore, offers an insight into pulser operation.

For the input stage, the inductance ratio is $L_{s1}/(2L_0)=G_1W_m/M_1$, where L_0 is the input linear inductor. Depending on the type of power supply used, G_1 may have a value slightly different than the G's for the other stages. However, for this simplified explanation, it is assumed that $G_1=G$.

Then the inductance ratio for all of the stages is

$$\frac{L_{sn}}{2L_0}=\frac{(GW_m)^n}{\Pi_1^nM_k} \tag{4}$$

This demonstrates that the core volume is directly proportional to the energy per pulse. For a given pulser, the pulse compression ratio ρ is fixed. ρ is defined as

---

Paper 55-179, recommended by the AIEE Special Communications Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 19, 1954; made available for printing December 16, 1954.

R. A. Mathias and E. M. Williams are with the Carnegie Institute of Technology, Pittsburgh, Pa.

Work reported in this paper was supported in part by the Rome Air Development Center under Contract *AF30(602)-96*.

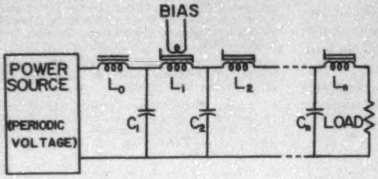


Fig. 1 (left). Basic magnetic pulser circuit

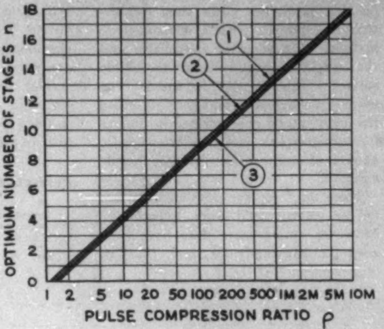


Fig. 2 (right). Optimum number of stages

Type of input:
1. Switched d-c, or half-cycle a-c per pulse with L_0C_1 tuned for maximum E_{m1}
2. Whole cycle a-c per pulse with L_0C_1 tuned to the a-c
3. Half-cycle a-c per pulse with L_0C_1 tuned to the a-c

$$\rho \equiv \frac{T_0/2}{t_{sn}} = \frac{\pi}{\omega_0 t_{sn}} \tag{5}$$

where

$$\omega_0 = \frac{1}{\sqrt{L_0C_1}}$$

$$t_{sn} = \left(\frac{e}{2}\right)^2 \sqrt{C_nL_{sn}} = \text{output pulse length}$$

Thus

$$\rho = \left(\frac{2}{e}\right)^2 \pi \sqrt{\frac{L_0}{L_{sn}}\,\frac{C_1}{C_n}} \tag{6}$$

But in the ideal case considered, $C_1 = C_2 = \ldots = C_n$. So the pulse compression ratio is

$$\rho = \left(\frac{2}{e}\right)^2 \pi \sqrt{\frac{L_0}{L_{sn}}} \tag{7}$$

This shows that, when the compression ratio is fixed, the over-all inductance and impedance ratio is also fixed and, as a result, the product of the core volumes is a constant

$$\Pi_1^n M_k = \text{constant}$$

From this it is noted that to minimize the total volume of core material to be required in the n stages, the volumes of all the cores should be equal. Or

$$M_1 = M_2 = \ldots = M_n \tag{8}$$

It is also of interest to find the number of stages required for the least total amount of core material for a given pulse compression ratio and energy per pulse. The total core volume is

$$M_T = M_1 + M_2 + \ldots + M_n \tag{9}$$

If this is minimized with respect to the number of stages n

$$n = ln\left[\rho^2\left(\frac{e}{2}\right)^4 \frac{2}{\pi^2}\right] = ln\left(\frac{2L_0}{L_{sn}}\right) \tag{10}$$

When the individual core volumes are equal

$$\frac{2L_0}{L_{s1}} = \frac{L_{s1}}{L_{s2}} = \frac{L_{s2}}{L_{s3}} = \ldots = \frac{L_{sn-1}}{L_{sn}} \tag{11}$$

So

$$\frac{2L_0}{L_{sn}} = \left(\frac{L_{sk-1}}{L_{sk}}\right) \tag{12}$$

From equations 10 and 12, it can be seen that

$$\frac{L_{sk-1}}{L_{sk}} = e \tag{13}$$

in which e is the Naperian base.

Equation 10 gives the optimum number of stages for the minimum total core material, and equation 13 shows that for these conditions the stage-to-stage inductance ratio is 2.72. Fig. 2 shows a plot of the optimums number of stages n as a function of the pulse compression ratio. Using more stages than this optimum number n is not economical, since losses are higher and more circuit components are required. In fact, the optimum design would be expected to involve somewhat fewer than the optimum number of ideal stages given in Fig. 2.

This conclusion arises when other considerations such as winding insulation and number of turns, iron and copper losses, and relative expense of associated capacitors are considered. Using 50–50 nickel-iron grain-oriented core material in a gapless type of construction, the value of G is about 1.0 in centimeter-gram-second units if the winding is close to the core. In this case, if the (ideal base) optimum number of stages are used, the least core material required is about 0.05 pound per stage per watt-second of pulse energy. The amount of core material per stage increases as the number of stages are reduced from the optimum value. However, in the case of the first stage core, the allowable copper losses and build-up of the winding may dictate a much larger than optimum size core if the volt-seconds integral to be absorbed by the main winding is large. This situation usually arises if the input is operating at power frequencies. A simple analysis of this first stage shows that the weight of core material per watt-second of pulse energy is directly proportional to the product of the input voltage and the core window area. Hence, a window of sufficient size in the first core may mean that the core volume has to be considerably larger than optimum size.

Iron losses, primarily because of the heat dissipation problems, may require

Nomenclature

t_k = time required to discharge kth capacitor
L_{sk} = saturated inductance of kth pulsactor
C_k = capacitance of kth capacitor, farads
ΔB_k = flux swing of kth pulsactor, gausses
E_{mk} = peak voltage on kth capacitor
N_k = number of turns on main winding of kth pulsactor
A_k = effective cross section of kth pulsactor core, square centimeters
μ_{sk} = saturated permeability of cross section enclosed by main winding, gausses per oersted; includes effect of core space factor, core case, insulation, and winding build-up
l_k = mean length of magnetic path of kth pulsactor, centimeters
M_k = volume of kth pulsactor, cubic centimeters
M_T = total volume of all saturable cores in pulser, cubic centimeters
W_{mk} = peak energy stored in kth capacitor, watt-seconds
G_k = a constant depending on geometry and core material of the kth pulsactor. The value usually is between 1.0 to 2.0 for Orthonol-type cores
ρ = pulse compression ratio, which is the ratio of 1/2 the period of the input stage to the output pulse duration
T_0 = period of input stage, seconds
t_{sn} = output pulse length, seconds
n = number of pulse compressing stages in the pulser

that the last few stages have smaller core volumes or special heat dissipation means. Another requirement is that, for maximum transfer of energy from one capacitor to the next, the capacitors increase in value inversely as the stage efficiency. This also implies that the peak voltage from one capacitor to the next drops in direct proportion to the stage efficiency.

Using the design principles outlined here, a magnetic pulser may be designed and the actual operating voltages and currents can be predicted with moderate accuracy. In some cases, transformer stages may be required to transform voltage level or impedance level since, when the input and output impedance levels are fixed, the compression ratio fixes the impedance ratio of the two unless linear interstage or saturable transformers are utilized.

Reference

1. THE USE OF SATURABLE REACTORS AS DISCHARGE DEVICES FOR PULSE GENERATORS, W. S. Melville. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 98, pt. 3, 1951, pp. 185–207.

Telephone Lines for Rural Subscriber Service

LESTER HOCHGRAF
NONMEMBER AIEE

R. G. WATLING
ASSOCIATE MEMBER AIEE

Synopsis: The cost of supplying telephone service to rural subscribers is high in large part because of the high initial cost and the maintenance of outside plant required to serve the subscribers. For the immediate future the most promising methods of reducing these costs are by new and cheaper methods of line construction and by concentrating the load on fewer pairs rather than by other types of transmission such as radio. Some Bell system developments now under trial to reduce the cost of line construction and to increase the load capacity of existing lines are described. The problems of carrier transmission in rural plant are discussed with emphasis on overall cost reductions which may be obtained by the use of the transistor and the compandor circuit.

IN THE rural areas they serve, the Bell system companies for many years have had as a major objective making phone service available at a reasonable price to every resident. Since the war over 2,250,000 Bell telephones have been added in rural areas. In doing so the companies have:

1. More than doubled the number of telephones in Bell rural areas.
2. Added telephones at the rate of over 1,000 every working day.
3. Added switchboards in many places and constructed many new buildings.
4. Spent around $750,000,000 to extend and improve service.

In addition to extending lines to families with no telephone service, considerable progress also has been made in improving the quality of rural telephone service. To help make such good progress possible, much development work has been done on new materials, methods and tools.

Open-wire and cable pairs are the types of facilities generally used to supply telephone service to rural subscribers in sparsely settled areas. Other means of transmission have failed to gain wide use because of the high initial cost of the equipment, the necessity for skilled maintenance, or because of expensive power supply arrangements needed to provide a dependable source of electric energy to the terminal equipment. Work is continuing in other fields, but for the immediate future it appears more promising to consider new and cheaper methods of building rural lines as well as the use of equipment, such as carrier systems, which will pemit more extensive utilization of outside plant facilities.

Probably no one method of construction nor any single new development will yield a large reduction in rural plant costs, but there are situations where different types of construction or new transmission arrangements can, and will, show cost savings. If advantage is taken of all such situations, the aggregate saving can be large and can make possible further rural plant extensions to serve new applicants for telephone service. Such a program is not glamorous, but it is necessary and is a challenge to the best brains that the industry can supply.

The rural telephone layout resembles a water shed map. Beginning at a small central office, nowadays usually a dial office, cables may radiate and serve nearby subscribers. As these cables become smaller in size, there comes a point where it is more economical to construct an open-wire line. Such a line might start with two or more crossarms of wire and taper in size as branches and taps serve subscribers. Finally, at the ends of the network, two and even single pair lines serve the more remote subscribers. Such a layout must include not only the distribution plant but the trunk circuits which tie the office to the system network.

At the close of World War II, the rural plant was made up largely of steel wire on pole lines having span lengths up to about 250 feet in heavy storm-loading areas. Steel line wire of the *109* size, good for these span lengths, had been made available in 1934. In 1937, a higher strength steel line wire *109H* was developed and span lengths of 350 to 450 feet were then practicable, the length of span depending on the storm-loading area. Zinc coatings applied by an electrolytic process which permitted coating weights two or three times those of hot dip coatings came into use in 1934, with consequent increase in life expectancy of steel line wire.

Although steel line wire provides a satisfactory transmission equivalent for the average rural line, a few of the open-wire pairs on some rural leads cover the longer than average circuits and for these a wire of higher conductivity must be provided. Copper is satisfactory from the transmission standpoint and is used where necessary, but its size and cost become prohibitive when it is made strong enough for long span construction. To care for this, copper-steel line wire (conductivity 40 per cent of that of the same diameter in copper) is used.

In 1946, extra high-strength steel line wire *109E* and a larger diameter steel wire *134H* of lower resistance were introduced. Both of these wires are good for maximum span lengths of 600 feet in heavy storm-loading areas, as contrasted

---

Paper 55-229, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted August 2, 1954; made available for printing December 8, 1954.

LESTER HOCHGRAF is with the Bell Telephone Laboratories, Inc., New York, N. Y., and R. G. WATLING is with the Bell Telephone Laboratories, Inc., Murray Hill, N. J.

to 350 feet for *109H* and 250 feet for *109*. Also, in 1947, a higher strength copper-steel wire (for use in spans up to 350 feet) was made available. This wire can be used when lower resistance is needed for transmission or signaling reasons.

The 600-foot span, which *109E* and *134H* wires make possible, is used by the Bell system operating companies not only in joint-usage pole lines carrying both telephone and power circuits but also in pole lines carrying telephone circuits only. It is pertinent to note that each pole eliminated from a line by using long span construction may represent savings of the order of \$30 to \$40 in first cost. The difference in wire costs is relatively small, about \$3 and \$20 per circuit-mile more for *109H* and *109E* respectively, as compared to plain *109* steel wire. The costs of copper-steel and of copper wire are approximately twice that of steel, but where transmission requires a facility better than steel, the use of copper-steel wire with long span construction results in a considerable saving in construction costs over that of solid copper wire.

Typical lengths of rural lines in various sections of the country range from 8 to 12 miles and in some instances extend as far as 25 to 35 miles in very sparsely settled areas. Such lines are not exempt from maintenance and transmission difficulties. The transmission and signaling qualities of long circuits are sometimes impaired during periods of rain, fog, and sleet because of insulator leakage. On the older rural lines, use of transpositions may be infrequent, with consequent prevalence of high circuit noise which has occurred as rural power networks have been extended. To control circuit noise and crosstalk when two or more pairs exist on a line, it is necessary that the circuits be transposed at certain intervals and in a prescribed manner. To be effective, the transpositions must be properly installed; such correct installation costs both time and effort.

Transposition Arrangements

Several transposition patterns have been developed over the years, such as the *ABC* system, the exposed line system, the *N* system, and the more recent *R1* system. The earlier systems employed "drop bracket" transposition assemblies to provide sufficient clearance between the wires when the plane of the wires was turned through 180 degrees. These transpositions require two spans to complete. The configuration causes a crosstalk irregularity which is appreciable when transpositions are closely spaced. A typical drop-bracket transposition is shown in Fig. 1.

The *R1* transposition system makes use of a tandem transposition bracket of the type shown in Fig. 2 or other "flat" type of transposition hardware. This provides a point-type transposition that can be "thrown in" easily after the wires are tensioned, thus providing some economies in construction. The *R1* design is characterized by having a transposition in each pair on alternate odd, or even, poles. This simplified design is relatively easy to administer in the field. The *R1* system is the present standard for rural line construction in areas relatively free from fog and sleet. It has given great improvement with respect to circuit noise as compared to the older designs. However, some service troubles may be experienced on tandem transposed lines in areas subjected to high winds and excessive precipitation. Under these adverse conditions, the use of a newly developed 4-inch point transposition assembly is recommended. This assembly retains the good features of tandem construction and provides a reduction of as much as 10 to 1 in leakage and greater freedom from wires contacting.

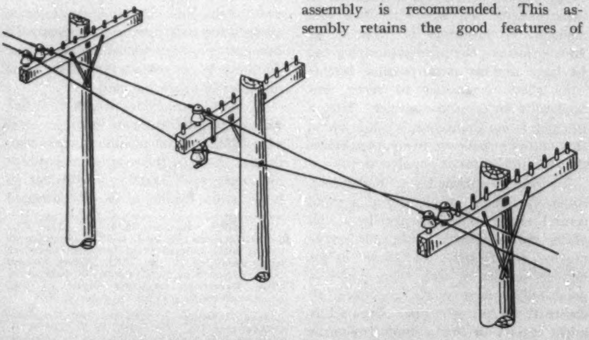


Fig. 1. Drop-bracket transposition

Rural Distribution Wire

Recently, a 6-pair wire facility has been developed for use principally in rural areas. This wire, known as rural distribution wire, finds its principal advantage in low material and construction costs and is intended primarily for use in reinforcing open-wire lines, particularly in locations where the addition of another crossarm would require pole replacements. The wire consists of six twisted pairs of individually insulated 19-gauge copper conductors, stranded around a polyethylene insulated *109E* steel line wire which serves as a support member. The copper conductors are insulated with high molecular weight followed by a jacket of polyethylene polyvinyl chloride compound which offers better resistance to abrasion than does polyethylene alone. The polyvinyl chloride compound is colored to provide ring, tip, and pair identification. Fig. 3 shows the method of fastening the rural distribution wire in a special pole-mounted bracket.

Suspension-Type Line

More recently attention has been directed by the Bell Telephone Laboratories to the development of a light open-wire line especially for use in sparsely settled areas. This line, designated the "suspension" type, employs the lightest materials consistent with good construction practice, and eliminates or simplifies many labor operations. Features of the new line are shown in Figs. 4 through 8. Substantially, all work incident to line construction is performed from the ground, thus virtually eliminating the necessity for pole climbing during line construction. Poles are equipped with pole brackets or crossarms, insulators, transposition brackets, and guy attachments before they are erected. Crossarms are bored to carry two or four pairs of wires. The suspension insulators are mounted underneath the crossarm, except at transposition positions, and are so designed that wires can be installed in them from the ground with a special lifting tool. Before the wires are raised to their insulators, vibration dampers are placed on each wire in each span. The damper is a short length of plastic tubing split helically along its length to facilitate its installation on the wire. This damper, by its flexing action as the wire moves under wind-induced vibrations, will absorb energy as fast as the wind puts energy in, thus damping out high-frequency low-amplitude vibrations which might result in injury to the wire at the point of its support. The damper is effective at any position on the wire in the span.

The wires are transposed by means of a simple repetitive transposition scheme. The line has been designed to provide up to four circuits at voice frequencies, which affords sufficient capacity to serve many rural areas. In those cases where more circuits are required, carrier systems could be used on some of the pairs. This line is still in the field trial stage, and it will be some time before advantages of the new system can be evaluated with certainty.

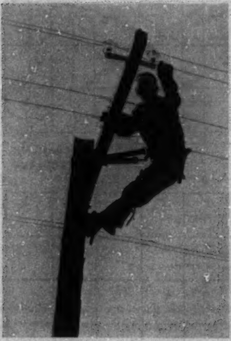


Fig. 2. Tandem transposition bracket on R1 line

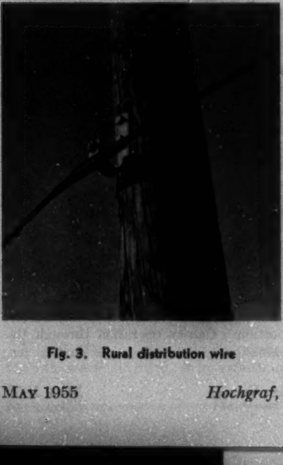


Fig. 3. Rural distribution wire

Carrier Possibilities

Two general means of concentrating traffic in the wire plant have been proposed and are under study. One is to obtain additional channels by frequency or time division means (carrier) and the other is to move switching devices closer to the subscribers and to give each subscriber access to more than one central office circuit.

Carrier possibilities for rural use have improved in the last few years by the development of the compandor, a circuit arrangement which provides a degree of immunity against noise and crosstalk, and by the development and availability of the transistor which offers exciting possibilities for reducing size of equipment and power supply requirements.

Subscriber carrier channels have already had some use in the Bell system. These channels were obtained by means of the type-*M* carrier system which provided one to five channels on a power line. Later modifications made it possible to use the *M* system on telephone pairs, but this use was limited by crosstalk to five channels per pole line. At present, about 10,000 rural subscribers are served by *M* carrier systems, some operating on telephone and some on power line facilities.

If it develops that carrier costs are reduced so that carrier channels can be used extensively, lines can be built with fewer pairs; fewer pairs can mean lighter line construction resulting in cost savings above those caused by the actual reduction in number of wires. Carrier channels would also be of particular value as a means for obtaining additional circuits on an existing line where it may not be possible to add more pairs without rebuilding the line.

It is not a simple problem to design, build, and operate a rural carrier system which can show substantial savings over wire distribution methods, even for the cases most favorable for carrier use. Such cases would be where the subscriber load is relatively concentrated and where the hauls are at least 4 or 5 miles long. Carrier transmission may be of particular value where it can defer major construction which would be required if voice frequency transmission were used.

Part of the high cost of present carrier systems suitable for rural use is the necessity for providing dependable power to operate the vacuum tubes in the terminal equipment. Where the carrier terminals can be located in telephone offices, this is not a severe problem but where, as in a rural subscriber carrier system, the carrier terminals must be remote from the office, this power supply problem may be a deciding factor. With the transistor now available, the hope is that the power requirements can be reduced to very modest values. It may be possible to supply stand-by power to transistors by cheaper methods than are possible where present vacuum tubes are used.

Fig. 4. Mounting crossarm on pole for suspension-type line

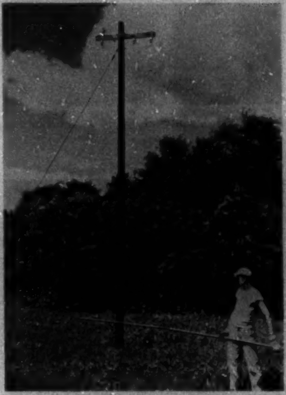


Fig. 5. Pole erected on suspension-type line

Transmission Characteristics of Rural Line

Possibilities of using carrier channels on existing rural plant vary from line to line. The extensive use of steel wire is an unfavorable factor because the carrier attenuation of steel wire pairs is relatively high. It is about ten times higher than for copper pairs which are used for toll carrier systems. Fig. 9 shows losses for *109* steel wire pairs and for copper-steel

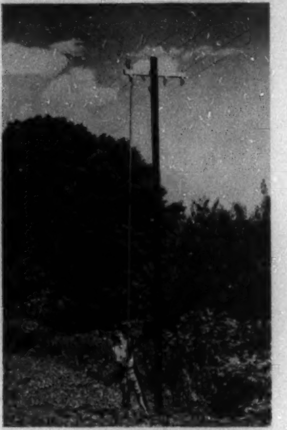


Fig. 6. Tightening of insulators on suspension-type line

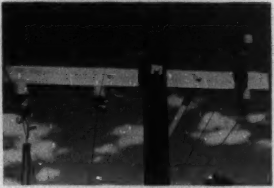


Fig. 7. Crossarm of suspension-type line with wires in place and being placed

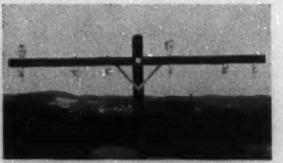


Fig. 8. View of suspension-type line with all wires in place

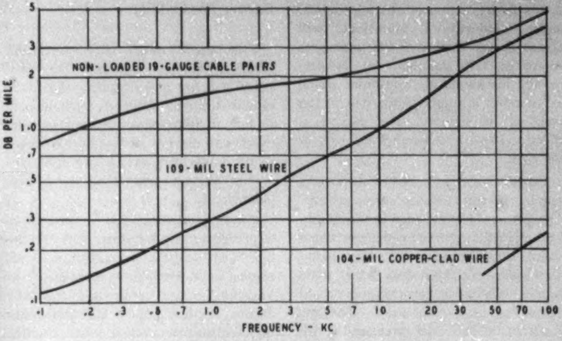


Fig. 9. Attenuation of typical wire circuits for rural carrier

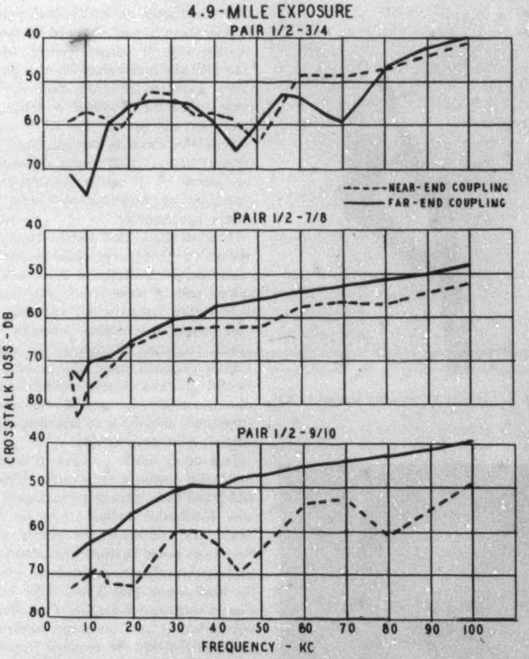


Fig. 10. Typical measured crosstalk couplings between R1 transposed pairs

pairs. A carrier system can be built to take care of line losses up to 30 to 40 decibels without repeaters. Carrier circuits on steel wire can be utilized for reasonable distances either by limiting the frequency band to that just above voice where attenuation is small, or by using repeaters or line amplifiers which, unfortunately, are themselves power-consuming devices. Where copper or copper-steel wire is used for voice transmission reasons, the carrier loss is low. On such pairs, the open-wire loss at carrier frequencies is not a significant factor for the distances involved in most rural circuits. However, most rural lines pass through a section of cable on the way to the central office.

It would, of course, be quite desirable to operate the cable portion of the circuits on a carrier basis as well as the open-wire portion. Carrier operation through the cable permits locating one carrier ter-

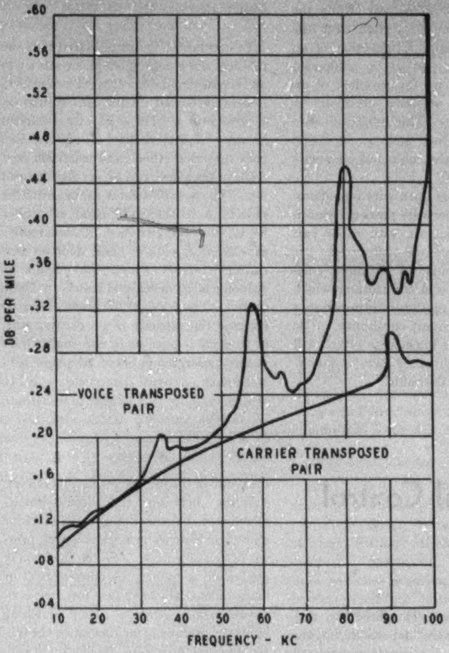


Fig. 11. Absorption losses of voice and carrier transposed pairs

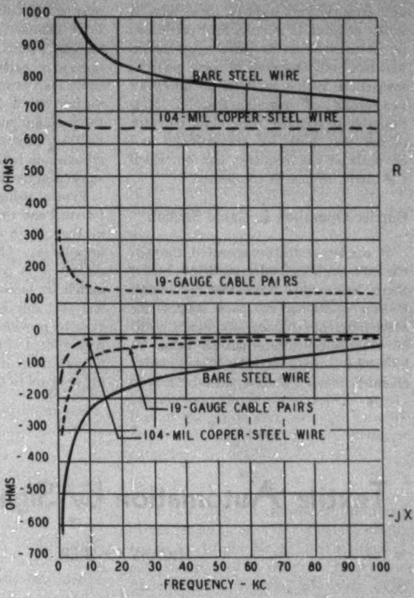


Fig. 12. Characteristic impedance of open-wire and cable pairs

minal at, or near, the central office and saves cable pairs as well as open-wire pairs. The carrier attenuation of cable pairs is quite high, as shown in Fig. 9. If the cable section is long, the result may be a carrier system in which the major part of the transmission loss is in the cable section, while the major savings are to be found in the open-wire portion of the circuit.

A favorable frequency band for carrier development is in the range from about 8 to 150 kc. The lower frequency is limited by the need for adequate frequency separation from the voice subscribers. The upper end is limited by the increase in attenuation, crosstalk, and possible interference from power line carrier systems and radio circuits. However, using this frequency band requires co-ordination with other carrier telephone systems operating on the same lines and using the same band of frequencies.

Crosstalk Considerations

It will often be desirable to operate more than a single rural carrier system on a line. In fact, it would be quite desirable to be able to apply carrier systems to any or all pairs on a line. Such operation is limited by crosstalk coupling between pairs. On most lines the pairs have been transposed only for voice operation.

Without compandors the crosstalk loss between line pairs should be held to the order of 60 decibels. With the voice transpositions, this value can be obtained only for frequencies just above the voice band or between those pairs that are well separated, as from one side of the pole to the other.

Compandors give a 20- to 25-decibel advantage against line crosstalk coupling. Where they are used, the coupling requirements are considerably relaxed and there is a great deal more freedom in the possible application of carrier channels to rural lines. Thus, the field for application of rural carrier will be enlarged by the development of a cheap compandor.

The rather general installation of *R1* transpositions in new rural plant is favorable to carrier development. A well-constructed *R1* transposed line can be counted on to provide two pairs on the top crossarm with low crosstalk coupling for carrier use up to 100 kilocycles per second on circuits up to 10 miles long with compandors. On any particular line, more pairs than this may actually be usable. Where it is desired to be certain of more pairs for carrier use, or where longer distances are involved, recently developed simple modifications of the *R1* design can be used. These modifications substantially decrease the crosstalk coupling.

Typical values of crosstalk coupling measured on an existing *R1* transposed line of good quality are shown in Fig. 10. For some carrier systems now under study, which use compandors, the crosstalk coupling should be better than about 40 decibels. This requirement depends very much on the system characteristics and on the level differences which are permitted to exist between the same frequency channels of carrier systems operating on different pairs.

On older lines, transposed 20 or more years ago, conditions may not be so favorable for carrier use. Because of the wide variety of transposition designs and arrangements used on these lines, it is not possible to make a firm prediction about these types of pairs. Absorption loss has been observed in some pairs at frequencies as low as 20 or 25 kilocycles per second, as shown in Fig. 11. It is desirable to measure these older pairs for carrier use, both for transmission and crosstalk, in the frequency band in which the carrier system is to operate.

Carrier Operation in Cable Section

If carrier is to be operated through the cable section, as discussed previously, there are some additional problems. Carrier operation on pairs loaded for voice frequency transmission is not possible since such circuits cut off below 10 kilocycles per second. The carrier channels in such a case must be routed over a nonloaded cable pair, while the voice channel may be routed over the original loaded pair. Connection of an open-wire carrier pair to a nonloaded cable pair brings up the problem of reducing the high reflection occurring at the junction point. This reflection leads to transmission loss and to increased crosstalk in both the cable and open-wire sections of the line.

Typical values of open-wire impedance for steel and copper wire pairs are shown in Fig. 12 in contrast with cable pair impedance. Above 30 kilocycles per second it is possible to obtain a good match by means of a transformer which will probably be of the autotransformer type to provide circuit continuity. The requirement of the impedance match will not be stringent but some form of matching appears to be desirable.

Conclusion

In conclusion, it can be stated that the problem of furnishing telephone service to an increasing number of rural subscribers is being attacked by the Bell system on several fronts. These include a careful scrutiny and a review of present line construction methods and materials, and trials of lines designed especially for rural use. The possible field of use of switching devices to concentrate rural subscriber loads is also being studied. Concurrently, experimental work is being done on new carrier systems which may make it economical to develop channels by these means. These might be used either to increase the capacity of an existing line, or to make it possible to use smaller lines to serve a large number of telephone subscribers.

Textile Automation by Signal Control

LOUIS CASPER
MEMBER AIEE

IN THE field of electric control of automatized machines, much effort is being directed towards co-ordinating the thinking of engineers in the mechanical field of design[1] and in the electrical field of signal communication.[2] One of the problems dictates the evolution of a common vocabulary. Closely related to it is the matter of making available to engineers in each of the two fields the body of knowledge already available in the other. These twin problems are discussed for a specific, typical situation.

The experimental development to be described looks to the general control of textile machinery by means of ordinary 5-unit, binary-coded paper tape and standard associated printing-telegraph apparatus[3] which are applied with minor alterations. A successful experimental model of an embroidery machine has been found capable of responding to tape control at the average rate of 500 stitches per needle-minute—the color, direction, and length of each individual stitch having been put under anticipatory control during the brief time interval separating each stitch from its precedessor. The development also contributes by specific example to the notion of the automatic factory in which unit machines may be ganged together and remotely controlled for purposes of labor economy and maximum machine-hour utilization.

Objectives

Because in this adaptation, as in previous typesetting-machine[4] and computer[5] applications, the control equipment has been lifted bodily out of a telegraphic environment, the remoteness of the control of single or ganged machines may be anywhere—at the units themselves, in one corner of the factory work room, floors below, or from an office or studio miles away. The last-mentioned capability is derived from the reproducibility of 5-unit telegraph signals by standard methods of tape reperforation, now technically feasible throughout the world by wire, cable, and radio transmission.

Granted that extreme remoteness of control, however practicable, may for the moment be brushed aside by today's textile practitioners as being somewhat in the realm of the future, yet flexible electric coupling between the needle-shuttle-fabric-feed structures of a textile machine and their controls remains a useful, present design objective. So too the ideal of electrically switching together several or many identical machines for common, simultaneous operation. Here switching holds the promise of emancipating the textile manufacturer from the burden of providing each one of a battery of machines with individual card or link-pin controls and from his present inability to deploy his facilities quickly to meet an influx of orders with their urgency priorities.

The Textile Machine Approach

Even in their modern forms, both weaving looms and embroidery machines tend to be heavy, ponderous, and inflexible because of the rigidity of their systems of control pins, rods, gears, cams, and pantographs, which must be mechanically coupled to the functional parts of the machines. When the controls are electrical, a relatively few solenoids interposed in the gears and shafting between driving motors and the actuated needles and fabrics suffice to impart to the latter all necessary movements in lieu of mechanical ties. The sequencing of the solenoids is adaptable to reliable methods of electric machine switching with its advantages of flexibility and speed.

EMBROIDERY IN COLOR

Textile embroidery usually uses a woven fabric as a base upon which the design is added as an embellishment. The

---

Paper 55-174, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing December 22, 1954.

LOUIS CASPER is with the Electrotext Corporation, New York, N. Y.

The author wishes to express appreciation to J. J. Fossaner, R. B. Benedict, and L. D. W. Casper for design and construction work in connection with the development.

machine resembles the familiar sewing machine in its needle movement. The fabric held in vertical movable frames and receiving the embroidery design must be positively positioned in sequential presentation to horizontal needles and their associated shuttles, loaded with colored or textured thread or yarn.

In the experimental machine, provision has been made for the initial use of ten different colors of thread. A compact row of ten horizontal needles in the *z* axis backed up by their corresponding ten shuttles (the operation of which can be separately selected for control) are rack-powered by a single drive designed for a $2^1/_2$-inch needle throw.

FRAME CONTROLS

In a vertical plane of the *x* and *y* axes which is interposed between the horizontal row of needles in front and the row of shuttles in back, the piece, or roll, of base fabric upon which the embroidered design is to be created is affixed semi-permanently to a pair of lead-screw driven frames running on guide rollers. The outer frame moves horizontally, to and fro, along the *x* co-ordinates in controlled steps equal to the chosen length of stitches and carries the vertical frame with its drive motor and lead screw. The counterbalanced inner frame moves vertically, up or down, along the *y* coordinate in stitch-length steps.

CONTROL BOXES

The codification of frame controls to accomplish stitching in 16 directions is simplified by the mechanical design of the power trains to the lead screws. Control boxes of identical pattern are interposed between horizontal and vertical frame drive motors and their respective lead-screws. In each control box are clutches and gear trains which perform two functions in response to electromagnetic controls: The lead-screws' progression, or the reversal thereof, is controlled by two clutch-trip magnets, either one being code selected for each frame movement; and the lead-screws' progression can be halved or restored to normal by two shift solenoids which are code selected to couple together, until reshifted, two pairs of matched gears in tooth ratios of 2-to-1.

By these means, a codified control tape advances the frames one step at a time in either direction, horizontally and/or vertically, in full- or half-stitch increments, as desired. In combination, the vertical and horizontal codes also advance the design in either direction along any one of six diagonals by simultaneous shifting of the frames in full- or half-stitches along both the *x* and *y* co-ordinates.

Table I. Directions of Needle

Direction	Vertical Movement	Horizontal Movement
Idle	none	none
N	up, full	none
NNE	up, full	right, half
NE	up, full	right, full
ENE	up, half	right, full
E	none	right, full
ESE	down, half	right, full
Etc.	up or down full or half	right or left full or half

DIRECTION OF MOVEMENT

The directions of relative movement of any fixed needle with respect to the changing position of the inner, vertical frame may be conveniently thought of as 16 compass directions from a datum point, as shown in Table I.

UNISON POSITIONING

One or the other of the end needles must have a home position, say at a corner of the design, from which it will be stepped, without stitches, to the point where the design will be begun. Area limit switches guarding the four design borders facilitate initial homing in a corner: The insertion in the coded tape of more than enough control steps towards the edges to pass the two border limit switches defining a corner will place a gang of machines in unison positions for commencing a new design, no matter what their respective locations may have been after previous use.

NEEDLE CONTROLS

All ten needles are normally retracted and latched in an inoperative position until released individually by coded signal. When all needles are latched, the fabric can be stepped by operation of the frames into position under the needle next to be released by the code tape control. When so released, the needle and its associated shuttle movement will produce one stitch.

SEVERANCE OF THREADS

The needle and shuttle threads are, in response to the tape control signal, severed close to the fabric on both sides of the workpiece.

COLOR CHANGES

Severance of needle and shuttle threads without human intervention and without loss of time beyond that of a single stitch facilitates ready changes of color in design since, to resume stitching, it is only necessary to serve the work horizontally (without stitching) to a needle shuttle loaded with thread of a proper color and then to release it.

POSITION CHANGES

The frames will also, in effect, transfer a needle, after retraction and severance of thread, from its former position to a desired new position anywhere on the workpiece.

TYPES OF STITCH

The prototype machine uses a lock-stitch reciprocating shuttle which, as in the familiar home sewing machine, locks, with a second thread, the loop introduced through the fabric by the needle and prevents its withdrawal with the needle. It is beyond the province of this paper to list the many textile variants for the insertion of thread or yarn through material which may vary from fine silk to heavy carpeting.

MATERIALS

Thread usable in embroidery machines of this type may be of any standard weight or gauge; it may be made of synthetics, silk, cotton, wool, or mixtures, from single filaments to heavy multiple-twist. The backing of embroidery may range all the way from flimsy gauze to burlap and canvas and range in shape from squares to belts to made-up garments.

LENGTH OF STITCH

As noted in the foregoing, the length of stitch is under continuous 2-to-1 codification in the machine. In addition, it is also adjustable, by presetting, from 4 to 32 stitches per linear inch.

The Printing Telegraph Approach

Five levers, each capable of taking up either of two positions, can determine 2^5, or 32, different code control combinations. These 32 combinations are sufficient to select and actuate any one of a combination of 26 different needles and associated shuttles, retract and relatch them, sever all threads, and halve the normal length of stitches.

There is a striking similarity between such control requirements and those met by a conventional teletypewriter. In the latter case, the same 32 code combinations are available; there are 26 different alphabetic type bars to be actuated, one at a time. The remaining codes are used to control the movement of the typewriter carriage and shifts between upper and lower cases.

Analogies

It was the similarity of telegraph communication art to the textile control problem which first suggested the probability that available telegraph equipment, with only minor adaptive changes, might be applied directly to the control of embroidery machines. In both cases, e.g., long 5-unit tapes can be code punched on keyboards (the telegraph perforator); they can be stepped along by means of punched out feed holes, one control code (the telegraphic letter) at a time, scrutinized by five transverse, ganged sensing levers (the telegraph transmitters' tongues), bearing electric contacts for making circuits (telegraphic marking), and breaking them (telegraphic spacing signals).

The Distributor

In action, telegraphic distributors are not unlike those of automotive ignition, with revolving arms sweeping sequential segments. In telegraphy, where miles separate perforator keyboards from terminating teleprinters, two additional commutator segments must be added to the five signalling segments to maintain necessary start-stop synchronism between the distributor brushes at the sending and receiving ends of the wire. These extra two segments, total seven, are usefully retained in textile control: first, to mark the time and set the operating pace in stitches per minute; second, to schedule and synchronize the related electrical events which must precisely dovetail; and third, to shift control from the 7-segment signalling channel just described and hereafter called the *B* channel, to a similarly contrived *A* channel having seven more segments; hence, a total of 14 segments during each revolution of the brushes before the entire cycle repeats. In telegraphy, the *A* and *B* channels are alike in function as well as in arrangement, and they transmit two different telegrams. In the embroidery machine, the two channels are different in their functioning: channel *A* governs the operation of the frames; and channel *B* governs the operation of the needles.

Two-Channel Operation

The necessity for a second channel in textile operation arises partly from the fact that the *B* channel's 32 code combinations have been exhausted in controlling the needles and, beyond these 32 combinations, there exists a further requirement for codes to position the frames carrying the fabric. Neglecting variable lengths of stitch, which have already been provided for on the *B* channel, frame movements require nine additional code combinations (including one for no frame movement). Another channel *A* is therefore required. Since three levers (affording 2^3-8 codes) are too few, a minimum of four levers is required for the *A* channel. It simplifies design to employ five levers instead of four, making channel *A* identical in form (though not in its functioning) with channel *B*, and making it possible to use the existing 2-channel, start-stop telegraph apparatus with little change.

There is another reason why the 2-channel telegraph adaptation proved attractive in textile control. The needles must be idle while the fabric is being moved and the fabric must be held steady while the needles pierce it and are withdrawn. Alternate, i.e., 2-channel, operation is the answer. As the distributor brushes sweep the first seven segments, the frames are moved to a new position. As the brushes sweep the second seven segments, the needles and shuttles accomplish one stitch.

At only one point has it seemed expedient to have the textile application depart from 2-channel telegraph practice. In the latter, telegrams on *A* and *B* channels are independently perforated in separate tapes by separate operators and are fed into separate transmitters. In textile control, it was evident that a single tape, fed through a single transmitter, was mandatory to preserve co-ordination between frames and needles. Therefore, alternate switching of the tape codes takes place—first frame function, then needle function; frame, then needle—*A* channel, then *B* channel, etc., as the tape progresses through the transmitter, center hole by center hole. There are therefore two center holes per stitch; two equal control-time intervals per stitch (seven segments, five code units); and two coded controls (frame, needles) per stitch.

Integration of Functions and Controls

Circuitry

Impulses through segments on the frame control channel actuate directly the trip magnets in the lead-screw control gear trains. Impulses through segments on the needle-shuttle selection and control channel actuate a series of five multipoint, single-pole, double-throw relays arranged in a fan or tree circuit in such a way as to find 16 different paths for the applied potential to take to trip, retract, and otherwise to control ten needles. In the experimental model the first of the five single-pole, double-throw relays has one pole idle; if utilized, the number of paths through the circuit could be increased by 16, and the total number of needles (and thread color changes) could be increased to 26.

A comparative sequence of the usage of 14 segments on 2-channel start-stop distributors and of embroidery machine control is now given:

In 2-Channel Telegraphy

1. Start.

2, 3, 4, 5, 6. Channel-*A* intelligence. 32 combinations, levers 1 to 5.

7. Stop.

1′. Start.

2′, 3′, 4′, 5′, 6′. Channel-*B* intelligence. 32 combinations, levers 1′ to 5′.

7′. Stop.

In Embroidery Machine Control

1. Start. Tape rests with *A*-channel punched holes over levers.

2, 3, 4, 5. Strapped for control of frames. During interval, frame relay and special secondary tripping circuit transfer any information on:

> Lever 1 to move vertical frame up.
> Lever 2 to move vertical frame down.
> Lever 4 to move horizontal frame to right.
> Lever 5 to move horizontal frame to left.
> Lever 3, if operative, renders levers 1, 2, 4, and 5 inoperative. Used to blank frame channel. Possible combinations: 1, 2, 4, 5, 1–4, 1–5, 2–4, 2–5, 3. Combinations not used: 1–2, 4–5; three pulses and up; five spacings.

6. Not used.

7. Steps tape, placing *B*-channel punched holes over levers.

8. Not used.

9, 10, 11, 12. Strapped for control of needles. During interval, needle relay and special secondary tripping circuit initiate stitching cycle, after having set up a 16-branch fan circuit controlled by levers 1 to 5 through relays as follows:

1. Levers 1–2–3–4–5 to release needle shuttle 1.
2. Levers 1–2–3–4 to release needle shuttle 2.
3. Levers 1–2–3–5 to release needle shuttle 3.
4. Levers 1–2–3 to release needle shuttle 4.
5. Levers 1–2–4–5 to release needle shuttle 5.
6. Levers 1–2–4 to release needle shuttle 6.
7. Levers 1–2–5 to release needle shuttle 7.
8. Levers 1–2 to release needle shuttle 8.
9. Levers 1–3–4–5 to release needle shuttle 9.
10. Levers 1–3–4 to release needle shuttle 10.
11. Levers 1–3–5 to retract all needles and sever all threads.

Vo
Pt.
19

12. Lever 1 to initiate horizontal half-stitches.
13. Levers 1–5 to initiate horizontal full-stitches.
14. Levers 1–4 to initiate vertical half-stitches.
15. Levers 1–4–5 to initiate vertical full-stitches.
16. Levers 1–3 to blank needle channel.

Combinations not used: 16 combinations omitting lever 1.

13. Not used.

14. Steps tape, placing *A*-channel punched holes over levers, preparatory to next cycle, pulse 1, start.

Ganged Machines

A cord-patching and switching panel is provided so that any one transmitter-distributor, out of a number of such, may be manually plugged up to any selected one or a multiplicity of embroidery machines, thus affording production flexibility and capacity to meet manufacturing demands.

Maintenance Costs

It is a feature of electric control, as opposed to mechanical control, that the patching board, transmitters, distributors, and switching may be concentrated in a lint-free location as remote as necessary from the machines proper. Interpretive neon lamps introduced into the circuitry facilitate electric servicing of a battery of machines; and machine concentration reduces testing and regulating labor requirements.

Since the electric equipment has seen long service in the telegraphic art, low costs for electrical maintenance may be reliably forecast, even in advance of factory application. Labor saving should be experienced in the mechanical maintenance department by elimination of cam and other relatively complicated mechanisms. Because of their reduced size and complexity, electrically controlled machines should cost less, produce savings in floor space requirements, and should consume less power.

Manipulation

Embroidery design is primarily a work of art. In textile manufacture the artist has the additional economic obligation of catering to the capabilities of his machine. In achieving beauty he must also choose a means of utilizing time most efficiently in presenting his fabric to the proper needles.

The artist, therefore, carefully makes in color a cross-section paper layout of the routes his needles must take to execute his pattern. From this layout, via a code schedule sheet, a keyboard operator punches tape, translating colors and needle routes into coded control signals.

Punching Tape

The textile keyboard, while still of telegraph standard tape-perforator design is, of course, no longer tied to the Roman alphabet. Its key caps are replaced with new ones designating the frame and needle controls; and internal changes governing punch block pin selection are made to secure for the operator a natural left- and right-hand cadence in controlling needles and frames. The operator must insert the necessary controls for frame movement where no stitching occurs, or where the artist has designated backtracking or stitch overlaying for special texture effects.

Short Cuts

Often-repeated coded tape sequences, such as those which guide the frames so that a needle stitches outline circles of various diameters or inward or outward spirals for delineating rings, become subjects for the advance preparation of template tapes and for their insertion in the master tapes at proper points to save the artist's and operator's time. Both men would naturally develop other cooperative short cuts. The preparation of control tape, though the means by which this machine achieves its versatility, is an exacting detail of manufacture.

Playback

Future experience may show that the principles of playback may supersede the described division of work between artist and keyboard operator. Before playback, the artist would use a standard arythmic teleprinter keyboard to create the original design in thread or yarn directly on the fabric. The tape, for reuse and duplication, would be produced as a by-product record of his original design.

Duplication of Tapes

The perforator is capable of producing more than one tape at a time, and by means of reperforators a tape may be used to reproduce any number of copies of itself economically. Paper tape may be run through transmitters scores of times before wearing out. Parchment tape is available, having life running into hundreds of occasions of reuse.

Multiple Patterns

In certain pattern designs advantage is taken of the needle flexibility afforded by code selection to select more than one needle for simultaneous operation. This means that the same design, in the same or different colors, will be repeated in different areas of the fabric during a single series of movements of the frames, the disposition of the areas of simultaneous repeat designs being a function of the spacial fixed separation between two or more selected needles of proper colors.

Mirror Patterns

Since patterns often take the form of mirror images across x and/or y center lines, it is obvious that the tape used to embroider one half or quadrant of a design can be used to perforate tapes for the other half or quadrants by mere reversals of x and y polarity codifications. Electric circuits are easily arranged to accomplish such reversals, contributing to the ease of preparation of repeated, reversed, and inverted design elements in control tapes.

Allover Patterns

By and large, embroideries are not allover patterns, but leave considerable areas of the base fabric exposed upon completion. In general, allover patterns therefore fall within the province of the weaving loom. However, weaving looms are not as versatile in some ways as embroidery machines, so that the electrically controlled, multiple-colored embroidery machine may find itself, in certain situations, in competition with the loom for producing allover patterns and even for solid textures of one color.

Solid or variegated colors may be embroidered economically row by row to achieve certain fabric textures not so readily obtainable by weaving, such as the chenilles. The simultaneous use of as many needles as practicable is indicated for such work. Manually thrown switches have been included in the experimental development for ganging selected needles for simultaneous release, use, and retraction, to embroider areas of solid or near solid colors.

Conclusions

It may be concluded that an initial design has been justified which has used a 5-unit 2-channel system for valid reasons beyond the obvious one of that system's availability. One of those reasons is that American communication companies are equipped[4] without changes in their circuits or apparatus to reproduce 5-unit control tapes at a distance and, simultaneously, at as many branch factories as may be desired.

From the results obtained by experimental operation of a prototype machine, it may be concluded that automation designers will find a valuable point of departure in the use of electromechanical units already familiar to telegraph engineers; and that the latter might profitably extend their normal interest in the transmission of intelligence signals over wires to comprehend the new uses of signals for the control of mechanisms, in the textile industry among others.

References

1. CONTROLLING MACHINE TOOLS AUTOMATICALLY, **F. W. Cunningham.** *Mechanical Engineering*, New York, N. Y., vol. 76, June 1954, pp. 487–90.

2. SYSTEM ASPECTS AND TRENDS OF MODERN COMMUNICATION, **I. S. Coggeshall.** *IRE Convention Record*, Institute of Radio Engineers, New York, N. Y., vol. 2, pt. 8, 1954, pp. 51–62.

3. **L. Casper.** U. S. Patent No. 2,649,065, Aug. 18, 1953.

4. TELETYPESETTER EQUIPMENT IN THE BELL SYSTEM, **M. N. Smalley.** *Bell Laboratories Record*, New York, N. Y., vol. 31, May 1953, pp. 183–86.

5. DIAGNOSTIC PROGRAMS FOR THE ILLIAC, **D. J. Wheeler, J. E. Robertson.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 41, Oct. 1953, pp. 1320–25.

6. TELEPRINTER CODES: 7-UNIT VERSUS 7.42-UNIT, **F. W. Smith.** *Western Union Technical Review*, New York, N. Y., vol. 8, Oct. 1954, pp. 129–34.

7. THYRATRONS IN LOADED CABLE MULTIPLEX, **H. H. Haglund, A. W. Breyfogel.** *Ibid.*, vol. 3, Jan. 1949, pp. 34–42.

8. AUTOMATIC CALL RECORDING AND ACCOUNTING IN THE SATT SYSTEM, **J. E. Ostline.** *AIEE Transactions*, vol. 72, pt. I, March 1953, pp. 56–62.

Rapid Measurement of Impedance and Admittance

B. SALZBERG
NONMEMBER AIEE

J. W. MARINI
NONMEMBER AIEE

Synopsis: A method is described for direct, rapid, and simultaneous measurement of the real and quadrature components of a driving-point or transfer impedance and admittance over a wide and continuous range of frequency. Basically, the method employs a mixer which derives its signal and local oscillator voltages from the voltage across and current through the unknown impedance respectively. With the unknown driven by a sinusoidal constant current, the mixer d-c output is linearly related to the unknown resistance. If the phase difference between the two derived voltages is altered by 90 degrees, the mixer d-c output is linearly related to the unknown reactance. To measure the components of an admittance, the unknown is driven by a sinusoidal constant voltage and the signal and local oscillator voltages are interchanged. Apparatus using this method for measurement of impedance and admittance in the frequency range of 50 kc to 5 megacycles (mc) has been built. With suitable modifications of detail, the method appears to be usable from acoustic to microwave frequencies.

THE increasing tempo of development and application of linear networks of all kinds has created a pressing need for impedance- and admittance-measuring apparatus with which results can be obtained rapidly and directly over wide and continuous ranges of frequency. A method for satisfying this need is described in the present paper; progress in this work has been reported in the past.[1,2] The method is applicable, in principle, from low frequencies to the very highest frequencies. It provides for rapid, direct, and simultaneous measurement of the real and quadrature components of a driving-point or transfer impedance and admittance over a wide and continuous range of frequency. Furthermore, it is well suited for visual presentation of the measurements and, hence, it is useful not only for the laboratory but for pedagogical purposes as well. Apparatus based on this method, designed for automatic or manual measurement of the two components of an impedance and admittance in the frequency range 50 kc to 5 mc, will be described and results of typical measurements will be presented. (During the course of our work we learned of a similar method and apparatus developed abroad.[3] Werthen and Nilsson apparently were unaware of the earlier work of Levy,[4] which seems to have anticipated much of their development. There are significant differences between the present paper and the work described in references 3 and 4.)

The steady-state currents in the meshes of a linear network resulting from the application of a voltage $V \sin pt$ to the driving-point terminals of the network are of the form $I \sin(pt+\theta)$. The relation between the amplitude of the applied sinusoidal voltage and the amplitude and phase of any particular mesh current is expressed by a quantity termed "impedance" or, by its inverse, "admittance." If the current under consideration is in the same mesh as the applied voltage, these terms are usually prefixed by the designation "driving-point"; if the current is in a mesh different from that in which the applied voltage appears, the prefix "transfer" is used. The generic quantities impedance and admittance are complex numbers whose values are determined by the angular frequency p and the parameters of the network. Although meaningful only for an applied sinusoidal force, these quantities are of fundamental importance because the response of a network to a nonsinusoidal force can be completely determined from a knowledge of their frequency dependence. These terms are defined as follows

$$\text{Admittance} = Y = G + jB = (I/V)\cos\theta + j(I/V)\sin\theta \quad (1)$$

$$\text{Impedance} = Z = R + jX = (V/I)\cos\phi + j(V/I)\sin\phi \quad (2)$$

where

G = conductance
j = quadrature operator
B = susceptance
V = amplitude of the voltage
I = amplitude of the current
θ = phase difference of current relative to voltage
R = resistance
X = reactance
$\phi = -\theta$ = phase difference of voltage relative to current

It is apparent from equation 1 that, if V is kept constant as the source frequency is varied, the conductance is directly proportional to the real component of the current, and the susceptance is directly proportional to the quadrature component of the current. The two components of an admittance can be determined, therefore, by maintaining V constant and measuring the quantities $I\cos\theta$ and $I\sin\theta$. Correspondingly, it is apparent from equation 2 that the two components of an impedance can be de-

---

Paper 55-211, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted September 7, 1954; made available for printing January 6, 1955.

B. SALZBERG and J. W. MARINI are with the Naval Research Laboratory, Washington, D. C.

The authors wish to acknowledge the work of G. S. Kan who developed an earlier model of the apparatus. They thank also K. W. Bewig for his help on many phases of the development of the present model, and D. T. Phillips who designed the swept-frequency oscillator used with the apparatus.

termined by maintaining I constant as the source frequency is varied, and measuring the quantities $V \cos \phi$ and $V \sin \phi$.

The real component of the current or voltage can be measured directly by means of a wattmeter[3,4] which records $VI \cos \theta$. For example, if V is kept constant, the wattmeter reading is proportional to $I \cos \theta$. Alternatively, if I is kept constant, the wattmeter reading is proportional to $V \cos \theta$.

Similarly, the quadrature component of the current or voltage can be measured directly by means of a reactive-power meter. This is a wattmeter provided with a means for shifting the phase between applied voltage and current by 90 degrees, so that the device[3,4] records $VI \sin \theta$. Thus, if V is constant, the reading is proportional to $I \sin \theta$; if I is constant, the reading is proportional to $V \sin \theta$.

The device used in the present work for measuring directly the real and quadrature components of current or voltage is a mixer of the kind customarily used for frequency conversion in radio receivers. Such mixers are simpler and more practicable than a wattmeter, are readily available, and are operable over very wide frequency ranges, from low frequencies to the highest radio frequencies.

When a mixer is used for frequency conversion, it is supplied with a large local oscillator voltage $P \cos pt$ and a small signal voltage v of different frequency. The mixer output current contains a component, the intermediate-frequency output, whose amplitude is proportional to the amplitude of the signal voltage and whose frequency is the difference between the signal and local oscillator frequencies. In the present application, however, the voltages supplied to the mixer have the same frequency. In this case, the intermediate-frequency output reduces to a d-c term that is proportional to the product of the amplitude of the small signal voltage and the cosine of the phase difference between the two voltages supplied. This can be shown formally by expanding the mixer output current in a Fourier series in pt and a MacLaurin series in v.

$$
\begin{aligned}
i(P \cos pt, v) = {} & [A_0(P)+A_1(P) \cos pt + \\
& A_2(P) \cos 2pt + \ldots] + v[B_0(P) + \\
& B_1(P) \cos pt + B_2(P) \cos 2pt + \\
& \ldots] + v^2[C_0(P)+C_1(P) \cos pt + \\
& C_2(P) \cos 2pt + \ldots] + \ldots \qquad (3)
\end{aligned}
$$

This is a general expression for the output current of a nonlinear device driven by a cosinusoidal voltage $P \cos pt$ and a second applied voltage v. If, in particular, v equals $Q \cos (pt+\theta)$ and Q is small enough to permit powers of v higher than the first to be neglected, equation 3 shows that the d-c output of the mixer is $A_0(P)$ plus the term $[(1/2)B_1(P)] [Q \cos \theta]$. The latter term arises from the product $[B_1(P) \cos pt] [Q \cos (pt+\theta)]$. Since this product also produces the intermediate-frequency output of the device when v has a frequency different from p, $(1/2)B_1(P)$ can be identified as the conversion transconductance $g_c(P)$ of the device. The d-c output can therefore be written as

$$I = A_0(P) + g_c(P) Q \cos \theta \qquad (4)$$

where $A_0(P)$ is the d-c output when the mixer is driven by a reference voltage $P \cos pt$ alone, and $g_c(P) Q \cos \theta$ is the incremental d-c output when a small voltage $Q \cos (pt+\theta)$ of the same frequency is then applied. This incremental d-c output can be used to determine the real component of a current or voltage and can, therefore, also be used to measure conductance or resistance.

Altering the phase difference between the two alternating voltages by 90 degrees before applying them to the mixer input causes the incremental d-c output to be proportional to $Q \sin \theta$. This product can be used to determine the quadrature component of a current or voltage and, hence, the susceptance or reactance.

Method

MEASUREMENT OF ADMITTANCE

A sinusoidal voltage source of constant amplitude is connected to the driving-point terminals of the network under test. If the driving-point admittance is to be measured, a current-sampling resistance r is inserted in series with the source; if a transfer admittance is to be measured, the sampling resistance is inserted in series with the transfer mesh. To insure that the current is essentially undisturbed by the sampling resistance, the resistance is made very small relative to the impedance of the circuit in which it is inserted. The reference voltage $P \cos pt$ across the driving-point terminals is applied to the input of a mixer; the resultant d-c component of the mixer output current has some constant value. The smaller voltage drop $rI \cos (pt+\theta)$ across the sampling resistance is proportional to the current, driving-point, or transfer as the case may be. This voltage is also applied to the mixer input, with the result that the d-c component of the mixer output current is changed in accordance with equation 4 from its constant value $A_0(P)$ by an amount $g_c(P) rI \cos \theta$. The change, as will be seen from equation 1, is proportional to the conductance being measured. The recording device which indicates the change in mixer d-c output current can be calibrated directly in units of conductance by substituting an admittance of known conductance for the network under test.

To determine the susceptive component of the admittance being measured, the phase difference between the two alternating voltages is shifted by 90 degrees before the voltages are applied to the mixer input. The resultant change in the d-c component of the mixer output current is then $g_c(P) rI \sin \theta$. This, as will be recognized from equation 1, is proportional to the susceptance being measured. If the same calibration as before is used, the recorder will indicate susceptance directly. Alternatively, the recorder can be calibrated by substituting an admittance of known susceptance for the network being measured.

MEASUREMENT OF IMPEDANCE

To measure the components of an impedance, the method is modified as follows: the constant-voltage source is replaced by a constant-current source; the latter is connected in series with a sampling resistance to the driving-point terminals of the network to be tested. The voltage $rI \cos pt$ across the sampling resistance is in phase with and proportional to the applied current; it is the first of the two alternating voltages which are applied to the input of a mixer. The second voltage is obtained from the input terminals of the network if the driving-point impedance is to be measured, or from the output terminals of the network if the transfer impedance is to be measured. In either case, this voltage has the form $V \cos (pt+\phi)$ so that the resultant change in the d-c component of mixer output current is $g_c V \cos \phi$. As will be seen from equation 2, this change is proportional to the resistance being measured. The recorder which indicates the change in mixer d-c output current can be calibrated directly in units of resistance by substituting an impedance of known resistance for the network under test. The reactive component of the impedance is determined in the same way, except that the phase difference between the two alternating voltages is shifted by 90 degrees before they are applied to the mixer input.

SWEPT-FREQUENCY MEASUREMENTS

This method of measurement not only provides direct readings of driving-point or transfer admittance and impedance components but may also serve for automatic and simultaneous presentation of the values of real and quadrature components. If the frequency of the constant-amplitude source is varied continuously over a band, at a relatively slow rate, the recorder will indicate the measured value of a component as a continuous function of frequency. Furthermore, if two mixers are used simultaneously, as previously described, readings of the values of both components as continuous functions of the frequency are made available. These readings of each of the components can be displayed separately as functions of the frequency or, if desired, they can be presented together in the form of a polar graph or circle diagram plot.

EVALUATION OF ERRORS

It will be noted that the method requires a source of constant amplitude—constant voltage if an admittance is to be measured, or constant current if an impedance is to be measured. This requirement demands the use of some form of automatic regulation or compensation. But even with a well-regulated source, there is invariably some departure from constancy of amplitude. As a result, errors in measurement can occur. In principle, if the output of the electronic device used to make the measurement were uninfluenced by variations in the amplitude of the reference voltage, the error would be equal to the fractional deviation from constancy of the source amplitude. Normally, the source amplitude could easily be maintained sufficiently constant to keep this error well within the usual engineering tolerances. In practice, however, the desired output of available mixers depends upon the amplitudes of both input voltages. This results in the introduction of two additional kinds of error, both of which are caused by variations in the amplitude of the reference voltage.

Referring to equation 4, if the amplitude P of the reference voltage applied to the mixer input varies, the output d-c component or reference output $A_0(P)$ also varies. The deviation from constancy of reference output constitutes one kind of error—the reference error. In addition, the conversion transconductance $g_c(P)$ is no longer constant. This is the source of the other kind of error, denoted as slope error.

In most cases, $g_c(P)$ varies slowly with changes in P, so that the slope error is less than the fractional deviation of P and is not very important. Reference error, however, can be serious. The value of Q in equation 4 must be kept small to insure linearity of operation. Consequently, the term $g_c(P)\ Q \cos\theta$, which is the output to be measured, may be only a small percentage of $A_0(P)$. Accordingly, a small percentage change in the latter term can produce a large error. In fact, under some circumstances, the reference error can be so large that the method of measurement becomes inoperative.

Reference error can arise from another source. In practice, the reference input voltage, represented heretofore as $P \cos pt$, also may contain harmonics of the fundamental angular frequency. The reference output $A_0(P)$ depends not only on the amplitude P of the fundamental but also on the presence of such harmonics. If the wave form of the reference input changes, as it might, e.g., in a swept-frequency measurement, reference error from this source occurs.

METHOD FOR ELIMINATION OF REFERENCE ERROR

While reference error can be reduced by careful choice of the operating conditions of the mixer to minimize the dependence of $A_0(P)$ on P or by use of a differential arrangement of matched tubes to balance out changes in $A_0(P)$, the error from this term can be eliminated altogether as follows: the amplitude Q of the smaller of the two alternating voltages derived from the network under test is modulated at some convenient audio-frequency before the voltage is applied to the mixer. In the case of sinusoidal modulation of Q at an angular audio-frequency a the mixer output of equation 4 becomes

$$\begin{aligned} I &= A_0(P)+g_c(P)Q(1+m\sin at)\cos\theta \\ &= A_0(P)+g_c(P)Q\cos\theta+mg_c(P)Q\cos\theta\sin at \qquad (5) \end{aligned}$$

The last term on the right of equation 5 is an audio-frequency term whose amplitude is proportional to the incremental d-c output in equation 4. This amplitude, therefore, can be used instead of the latter output to measure the components of an admittance or impedance.

If a band-pass filter centered at the modulating frequency is used as the output circuit of the mixer, the output voltage is made unresponsive to all other components of the mixer current, eliminating thereby the large d-c reference output $A_0(P)$ and the reference error associated with it. A linear voltmeter used to measure the amplitude of the output voltage from the band-pass filter, however, would provide only a positive output. Consequently, to determine the signs of the components as well as their magnitudes, phase-sensitive demodulation[5] must be used.

MEASUREMENT OF THE QUADRATURE COMPONENT

As stated previously, measurement of the quadrature component requires a 90-degree alteration of the phase difference between the two alternating voltages which are derived from the network under test and applied to the mixer. If measurements are to be made manually at a single frequency, this requirement can be met quite simply, e.g., by processing the reference voltage through an resistance-capacitance circuit. However, if measurements are to be conducted automatically over a wide frequency range, the situation becomes more involved. Not only must a 90-degree phase shift be obtained throughout the entire frequency range but the amplitude of the reference voltage must also be maintained constant to within the limits imposed by the maximum allowable reference and slope error.

One method for obtaining the 90-degree phase shift is to heterodyne the two voltage samples obtained from the network under test to a fixed intermediate frequency before applying them to the mixer. The required phase shifting is then easily accomplished at the intermediate frequency. This method can be used not only at acoustic and video frequencies but also, with appropriate changes in the means for obtaining voltage samples, at microwave frequencies as well. The method, however, requires auxiliary converters and a tracking heterodyne source whose frequency differs from that of the main source by the intermediate frequency.[3,4]

In principle, the 90-degree phase-shifting requirement can be circumvented by using a source which provides two constant-amplitude outputs, the phase difference between these being 90 degrees at all frequencies.[6] One of the two outputs is used to drive the network under test and to provide a constant-amplitude reference voltage for measuring the real component. The other output is used as a 90-degree phase-shifted reference voltage of constant amplitude for measuring the quadrature component. However, if admittances of high values are to be measured, the driving output-voltage source should have a very low internal impedance so that the 90-degree phase difference is maintained between the voltage applied to the network and that supplied by the auxiliary output. (The latter requirement does not apply to the measurement of impedance.)

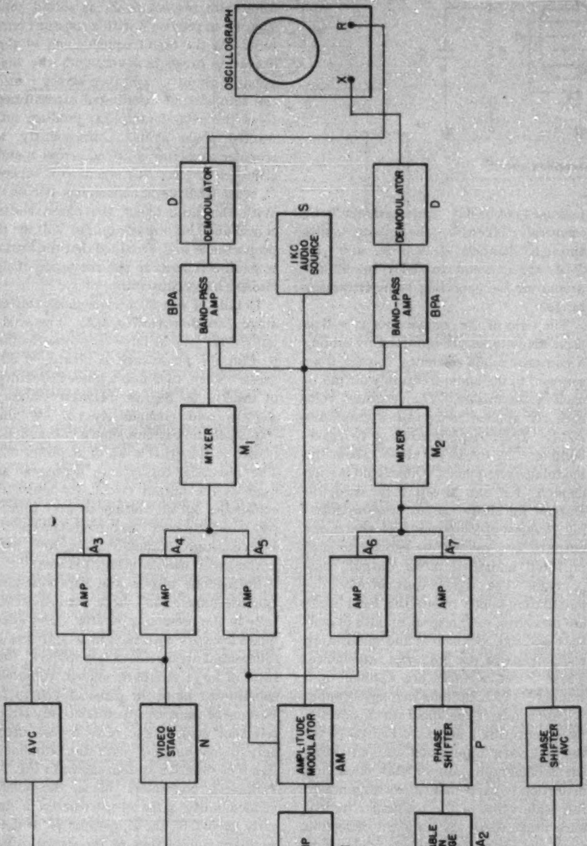

OSCILLOGRAPH
R
X
D
DEMODULATOR
S
1KC AUDIO SOURCE
DEMODULATOR
D
BPA
BAND-PASS AMP
BAND-PASS AMP
BPA
MIXER
M1
MIXER
M2
A3
A4
A5
A6
A7
AMP
AMP
AMP
AMP
AMP
AVC
VIDEO STAGE
N
AMPLITUDE MODULATOR
AM
PHASE SHIFTER
P
PHASE SHIFTER AVC
AMP
A1
VARIABLE GAIN STAGE
A2
r
Z
I
V
VARIABLE VIDEO FREQUENCY SOURCE

Fig. 1. Impedance-measuring apparatus, block diagram

A third method for shifting the relative phase of the two voltages by 90 degrees is the employment of a wide-band phase-shifting network.[7] A fourth method, employed in the apparatus described in the following, makes use of a simple resistance-capacitance circuit which provides a 90-degree phase shift together with an automatic volume control (avc) system which maintains the output constant as the frequency is varied.

Apparatus

Laboratory apparatus for automatic or manual measurement of the real and quadrature components of a driving-point or transfer impedance and admittance has been developed for the video-frequency range of 50 kc to 5 mc. The available impedance ranges are from 0 to 100 and 0 to 1,000 ohms. The available admittance ranges are from 0 to 100, 0 to 1,000, and 0 to 10,000 micromhos. Automatic measurements are made by sweeping the source frequency through convenient portions of the video range at a rate of 1 cycle per second or less, depending on the characteristics of the network being measured. The real and quadrature components are measured simultaneously and the results are displayed on a cathode-ray oscilloscope.

The apparatus used for measuring the driving-point impedance is shown schematically in Fig. 1. A variable video-frequency source of sinusoidal voltage, the amplitude of which is automatically controlled, is connected in series with a resistor r to the unknown impedance Z. The value of r is 100 times greater than the maximum value of the impedance range used; the current I through Z, therefore, is limited by r and is essentially independent of Z. The small voltage IZ developed across Z is used as a signal voltage. The relatively large source voltage V, essentially a replica of the voltage rI developed across r, is used as a reference voltage.

The signal voltage is amplified by a single-state video amplifier A_1 and then amplitude-modulated at 1 kc in grid-circuit modulator AM, where a modulating voltage of constant amplitude is supplied from audio source S. To suppress undesired audio-frequency modulation products, a high-pass filter is used for the output circuit of the modulator. The amplitude-modulated signal voltage is applied to two identical single-stage video amplifiers A_5 and A_6 followed by like parabolic-law mixers M_1 and M_2.

The reference voltage is applied through a resistance-capacitance video circuit N

1, 74
1
55

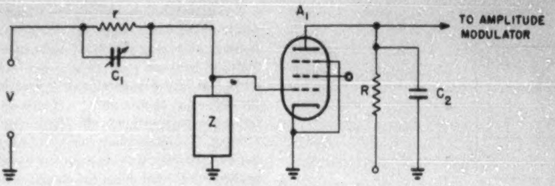


Fig. 2. Impedance-sampling circuit

to two identical single-stage video amplifiers A_3 and A_4. The output of A_3 is connected to an avc which controls the video source to maintain constant the amplitude of the reference voltage at the outputs of A_3 and A_4. The output of A_4 is connected to the input of M_1 through a plate load circuit common to both A_4 and A_5. Mixer M_1 is thus supplied with a small amplitude-modulated video signal voltage and a relatively large video reference voltage of constant amplitude.

The reference voltage is also applied to a variable-gain single-stage video amplifier A_2 which is followed by a 90-degree phase shifter P and a single-stage video amplifier A_7, whose output circuit is common to A_6. The reference voltage is applied through this common circuit to the input of M_2, and also to an avc which controls the variable-gain stage so as to maintain constant the amplitude of the phase-shifted reference voltage at the mixer input. Mixer M_2 is thus supplied with a small amplitude-modulated video signal voltage and a relatively large 90-degree phase-shifted video reference voltage of constant amplitude.

It is essential to preserve the relative phase between signal voltage IZ and reference voltage V; for this reason, the phase characteristics of the signal-voltage channel A_1–AM–A_6 and reference channel N–A_4 are matched over the operating video-frequency range. Because the avc which controls the video source is operated from the output of A_3 instead of directly from the source output, the response characteristics of these two channels need not be flat; they need merely be identical. Likewise, the phase shifts through channels A_1–AM–A_6 and A_2–P–A_7 are adjusted to differ by 90 degrees over the operating video-frequency range.

The signs of the components as well as their magnitudes are obtained by adding a constant audio-reference voltage from source S to the audio-frequency output of each of the mixers. The resultant voltages are applied to bandpass amplifiers BPA. These amplifiers have a flat passband of ±300 cycles per second about the modulating frequency. This band is wide enough, for the sweep rate used, to amplify faithfully the varying modulating frequency components of the mixer outputs, and yet narrow enough to suppress all undesired mixer output components. The d-c output of the demodulators which follow the BPA units are measures, with respect to both magnitude and sign, of the real and quadrature components of the unknown impedance. These direct voltages are applied to a standard 5-inch cathode-ray oscillograph.

Calibration of the apparatus is effected by substituting a standard resistance for the unknown impedance and adjusting the oscillograph gain so that maximum deflection in each channel corresponds to full scale value of the particular impedance range used. A standard deflection from the output of the channel which measures the quadrature component is obtained by switching out the phase shifter P during the calibrating process.

The circuit used to obtain a signal voltage from the unknown impedance is shown in Fig. 2. The resistor r is made large with respect to Z, as stated previously, to provice Z with a constant current. At the high-frequency end of the operating range however, both the distributed shunt capacitance across r and the total output distributed capacitance C_2 of the pentode amplifier produce unwanted phase shifts. Consequently, a trimmer capacitor is added across r and adjusted so that RC_2 equals rC_1, where C_1 is the total shunt capacitance across r. With this adjustment the phase shifts cancel and the pentode output voltage is proportional to Z, provided that the latter is small compared to the reactance of C_1 and the input capacitance of A_1.

Details of the 90-degree phase-shifter stage P are shown in Fig. 3(A). The product RC is made as large as is practicable so that the phase shift is nearly 90 degrees. Then, to obtain a phase difference of exactly 90 degrees between channel A_2P–A_7 and channel A_1–AM–A_6, the complementary circuit shown in Fig. 3(B) is used in the grid circuit of amplifier A_6. The resistance-capacitance arrangement used as the output circuit for stage P causes the output voltage to vary inversely with frequency; to provide the necessary constant output voltage, and avc arrangement shown in Fig. 1 is used.

Both diode mixers and parabolic-law pentode mixers have been used successfully in the present apparatus. For adequate linearity with such mixers, unless a differential arrangement is employed, the ratio of large reference voltage to small signal must be of the order of 100 to 1. Because of the consequent relatively large reference error, these mixers were used only in conjunction with the auxiliary signal modulation system shown in Fig. 1. Pentagrid converters alone, however, yielded fairly satisfactory results in an early model of the apparatus in which auxiliary modulation of the signal voltage was not employed. With such multigrid tubes, the output was linear for ratios of large reference voltage to small signal voltage as low as 10 to 1, provided

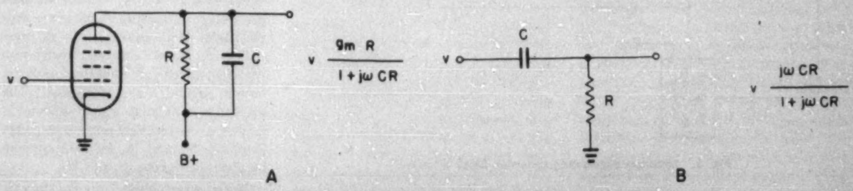


Fig. 3. (A)—Ninety-degree phase-shifting circuit. (B)—Compensating network

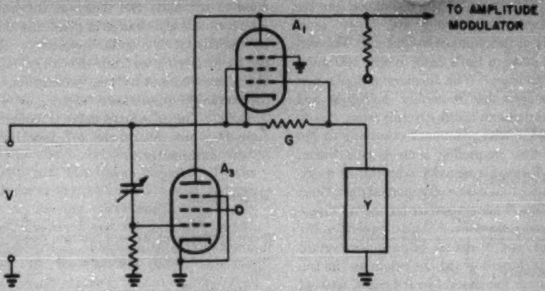


Fig. 4. Admittance-sampling circuit

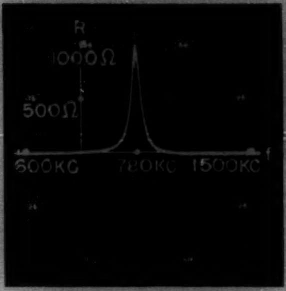


Fig. 6. Resistance versus frequency for a parallel resonant circuit

proper polarizing voltages were chosen. A large screen-grid resistor was used to minimize reference error resulting from changes in amplitude of the large reference voltage.

The video source is a heterodyne-type oscillator consisting of an electronically swept phase-shift oscillator,[8] variable from 20.050 to 25 mc, whose output is mixed in a frequency converter with a 20-mc signal from a fixed oscillator to produce a 50-kc to 5-mc output. The amplitude of the output is controlled by varying the gain of a tuned buffer amplifier between the fixed oscillator and the frequency converter.

To measure the components of a transfer impedance instead of a driving-point impedance, the video source is connected in series with resistor r to the driving-point terminals, as shown in Fig. 1, but the small signal voltage is now derived from the transfer mesh. The results of measurement thus represent real and quadrature components of the voltage in the transfer mesh, referred to the driving-point current. By definition, this is the transfer impedance.

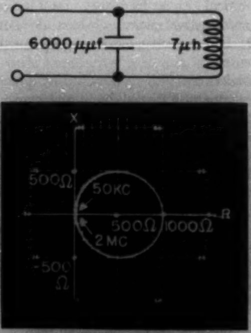


Fig. 5. Reactance versus resistance for a parallel resonant circuit

It is usually desirable to be able to ground one terminal of both the video source and an unknown impedance or admittance being measured. This can be done, when measuring an impedance, by deriving the reference voltage from the source, as shown in Fig. 1, rather than from the voltage developed across the current-limiting resistor r. When the driving-point admittance is measured, however, it is necessary to maintain a reference voltage of constant amplitude at the driving-point terminals and to derive the small signal voltage from a series-connected current-sampling resistance. Under these conditions, the sampling resistor must be floated above ground. Details of the arrangement actually used for measurement of an unknown driving-point admittance Y are shown in Fig. 4. The conductance G of the sampling resistor is large compared with Y; in practice, G is 100 times greater than the maximum value of the admittance range used so that the source voltage is essentially entirely across Y. A floating pentode amplifier A_1 is used to sample the voltage drop across G. A second pentode is used to balance out the effect of stray capacitance across Y and of screen-to-plate capacitance of A_1. If, in practice, G cannot be obtained as a pure conductance, the error otherwise introduced by the susceptive component can be eliminated by replacing the plate load resistor of the pentodes by a suitable compensating network.

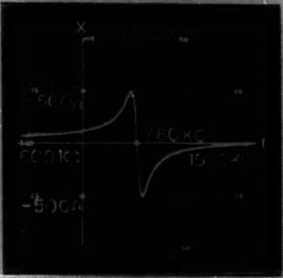


Fig. 7. Reactance versus frequency for a parallel resonant circuit

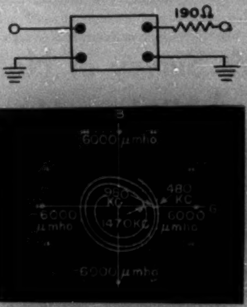


Fig. 8. Transfer susceptance versus transfer conductance for a delay line

To measure the components of a transfer admittance instead of a driving-point admittance, the video source is connected directly to the driving-point terminals. The voltage across these terminals is used as the large reference voltage, and the small signal voltage is derived across a series sampling resistance inserted in the transfer mesh.

Typical Results

Typical results obtained with the laboratory apparatus described in the foregoing indicate that the principles employed are sound and that design of useful apparatus of this type is feasible. The error of the apparatus in any measurement depended on the impedance and frequency ranges involved. Spot checks indicated that this error was usually within the range of linearity of the oscillograph used as a recorder, which was about 5 per cent. In some cases this error was exceeded, however, especially at low or high frequencies where the video channels failed to line up identically. Typical results obtained with the apparatus are shown in Figs. 5 through 8. These photographs are time exposures taken with a Polaroid-Land camera.

Figs. 5 through 7 resulted from the measurement of a parallel resonant circuit consisting of a 6,000-micromicrofarad capacitor in shunt with a 7-microhenry inductor which has a Q of 30 at the resonant frequency of the circuit. Fig. 5 shows a polar plot of the input impedance of the circuit over the range from 50 to 2,000 kc. The resistance and reactance can be read from the horizontal and vertical axes respectively. The scale is 100 ohms per division so that the maximum value of the equivalent series resistance is seen to be 1,000 ohms. This agrees with the value calculated by multiplying the Q, measured on a Q meter, by the reactance of one of the components at resonance $Q(L/C)^{1/2}=30\ (7\times10^{-6}/0.006\times10^{-6})^{1/2}=1{,}000$ ohms.

In Figs. 6 and 7 the resistance and the reactance respectively are shown separately as functions of frequency. The vertical scale in both cases is still 100 ohms per division. The horizontal scale now represents the frequency, however, and the range exhibited extends from 600 to 1,500 kc, with resonance occurring at 780 kc. The frequency scale is not linear. Fig. 8 shows a transfer admittance measurement taken on a commercial delay line. The line is advertised as having a "Gaussian response curve." Consequently, the tip of the transfer admittance vector spirals down toward the origin as the frequency is increased toward cutoff instead of making a circle about the origin as is the case for lines with the conventional flat response. The scale is 600 micromhos per division.

The frequency markers and the origin in Fig. 8 were inserted photographically by a method described by Levy.[4] The curve for the delay line was photographed with the video-frequency source sweeping automatically over the desired range. The automatic sweeping mechanism was turned off and the source frequency was adjusted by means of a calibrated dial to 480, 980, and 1,470 kc successively, an additional exposure of the photographic plate being made at each frequency. Finally the network was switched out and an exposure was made to locate the origin. The frequency markers can be used to calculate the delay of the line, which equals the change in phase divided by the change in angular frequency. As the frequency varies from 480 to 1,470 kc, the phase changes by two revolutions of the transfer admittance vector, or 4π radians. The calculated delay is 2 microseconds, which is also the delay specified by the manufacturer.

References

1. Rapid Measurement of Admittance at Radio Frequencies, Bernard Salzberg, George S. Kan. *1948 Joint Meeting of Institute of Radio Engineers and URSI*, Washington, D. C., Oct. 1948.

2. Measurement of Impedance and Admittance, Bernard Salzberg, J. W. Marini. *1952 Institute of Radio Engineers' National Convention*, New York, N. Y., March 4, 1952.

3. An Automatic Impedance Meter, H. Werthen, B. Nilsson. *Transactions*, Royal Institute of Technology, Stockholm, Sweden, no. 8, Jan. 1947, pp. 1095.

4. Methods and Apparatus for Measuring Phase Distortions, M. Levy. *Electrical Communications*, New York, N. Y., vol. 18, Jan. 1940, pp. 206–28.

5. Waveforms (book), Chance, Williams, Hughes, Sayre, MacNichol. McGraw-Hill Book Company, Inc., New York, N. Y., 1949, sect. 14.4–14.5.

6. Regeneration Theory and Experiment, E. Peterson, J. G. Kreer, L. A. Ware. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 22, Oct. 1934, pp. 1191–1210.

7. Synthesis of Wideband Two-Phase Networks, H. J. Orchard. *Wireless Engineer*, London, England, vol. 27, March 1950, pp. 72–81.

8. Reactance Tube Modulation of Phase Shift Oscillators, F. R. Dennis, E. P. Felch. *Bell System Technical Journal*, New York, N. Y., vol. 28, no. 4, Oct. 1949, pp. 601–07.

A Survey of Noncontacting Vibration Pickups Using Electric Fields

HOWARD F. CLARKE
NONMEMBER AIEE

THE growing importance of vibration requirements in the field of component testing has created an increasing need for means of measuring the magnitude of the vibrating forces and displacements to which components are subjected. In general, quantities should be measured; acceleration, velocity, and displacement. While it is always advantageous to measure directly the quantity to be known, it is quite easy to convert data from one quantity to the other when the frequency and wave form are known.

There is a large variety of methods, electrical, mechanical and optical, which may be used to measure vibrations; however, by far the most common method is to convert the vibration into electric signals, and then to measure the electric signal. If this procedure is followed, the main problem of the measurement lies in the transducer making the conversion.

There are a number of devices on the market for the conversion of vibrations to electric signals, which might be divided into three classes, according to whether displacement, velocity, or acceleration is the quantity intended to be measured. Displacement and velocity can be converted directly to electric quantities, and several transducers are available for this purpose. Acceleration is usually mechanically converted to displacement, using a spring-mass system operated well below resonance, and the displacement is then applied to a displacement transducer. It should be noted that this places theoretical upper-frequency limitations on acceleration pickups which do not hold for velocity and displacement pickups.

Acceleration may be obtained by direct measurement of displacement and knowledge of the wave form and frequency. There is a disadvantage to the use of displacement-measuring devices if acceleration constant over a considerable frequency range. Nevertheless, the freedom from the disadvantages of resonating systems is of sufficient importance

---

Paper 55-168, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing January 21, 1955.

Howard F. Clarke is with the Boeing Airplane Company, Seattle, Wash.

to make displacement-measuring systems of considerable use in vibration measurement. In this paper a class of displacement-measuring devices is described.

In the systems to be discussed, no mechanical connection is made to the vibrating body, the vibrating displacement being sensed by its effect on an electric field between the vibration pickup and the object shaken. In the transducer system shown in Fig. 1, there are no moving parts except the object being shaken, thereby eliminating pickup resonances or modes of vibration which arise to trouble the user. Another distinct advantage is that the pickup does not mechanically load the vibrating body. The system may be arranged to have negligible side-sensitivity and almost no practical limit to its frequency response from zero frequency up. This means that these devices are especially well adapted for use as transfer devices between static and dynamic displacements as they may be calibrated statically.

Perhaps the greatest disadvantage to electric field systems is the fact that their sensing field cannot be confined to a small area without also confining their linear range. A second disadvantage is that the vibrating surface which they see must usually have some known geometry, such as a flat sheet. Also, the nature of the material of the vibrating surface seen is critical, the systems being chiefly set up to see surfaces of nonferromagnetic materials of high conductivity. Their response to a surface of geometry and material other than the foregoing will be described as each method is taken in turn. The three methods of using electric fields to pick up vibrations will be discussed in detail in the following.

Capacitance-Type Vibration Systems

Historically, the first field-type pickup was the capacitor microphone. As this is quite familiar to everyone, it will not be discussed further except to consider it as a possible vibration pickup for the purpose of comparison with some of the systems to be discussed.

The chief disadvantage of capacitive vibration pickups would appear to be the difficulty in obtaining sensitivity and linearity with small probe size. In the capacitance-type pickup, the capacitance of the sensing element itself is often no greater than the leads connecting it to the following circuit. Since the capacitance of these leads is in parallel with the capacitance of the sensing element, the effective percentage change in the sensing-element capacitance, as seen by the circuit, becomes smaller as the leads grow larger.

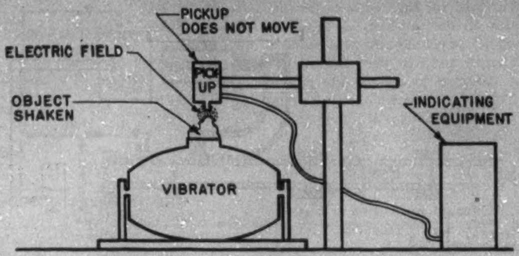


Fig. 1. Vibration setup using field pickup

The frequency response of the ordinary microphone will not extend down to direct current because of the use of d-c excitation through a shunt resistor, and, therefore, such a device will not serve as a transfer between static and vibratory displacements. For static to dynamic displacements, a carrier system must be used. The capacitance probe could be used to determine the frequency of a frequency-modulated system but, here again, problems of linearity may be encountered. For stability and noise level the capacitance probe appears comparable with other systems and could not be disturbed by ferromagnetic material in the target.

It is quite possible that methods will be found of overcoming or minimizing the shortcomings of such systems so that they may be considered as comparable in possibilities with the other electric field systems to be described. However, the foregoing reasons have tended to channel development away from capacitance-type pickups at Boeing Airplane Company.

Microwave Systems

These systems use assemblies of standard-type plumbing components in the region of 10,000 megacycles to pick up vibration by the changes in the power which is reflected back into a radiating system by the vibrating surface when it is placed close to the source of radiation. The aforementioned frequency is by no means suggested as optimum, but is rather a convenient high frequency for which a wide variety of components are commonly available.

Microwave systems are somewhat related to the capacitance-type pickup in which a carrier frequency is used. However, it has been pointed out that in a capacitance-type pickup, capacitance changes in a small sensing element tend to be lost in lead capacitance. With high frequencies, a change in even a very small capacitance appears as an appreciable susceptance change and, if this is applied at the termination of a transmission line, the resulting change in reflection coefficient of the termination, as seen at the input terminals of the line, is not appreciably diminished by increasing the length of the line. Thus, systems using very high frequencies have two advantages; a small change of geometry shows a comparatively large change in reactance or susceptance, and this change may be propagated down a transmission line without serious loss.

Of the three systems worked out by the author, two are sufficiently useful to warrant a description here.[1] In Fig. 2, a microwave oscillator 2 feeds energy to a wave guide 3 through a dissipative pad 4 which has as its purpose the minimization of disturbances in the oscillator out from

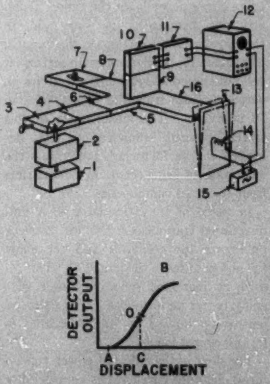


Fig. 2. Magic T field pickup

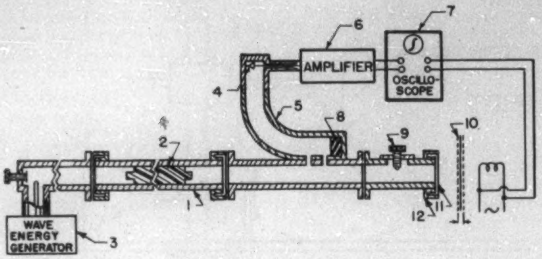


Fig. 3. **Directional coupler field pickup**

variations in the impedance of the load. From the dissipative pad, the power enters a branched guide of the type known as the magic T, 5. On one horizontal arm 8 of the magic T is placed an impedance, arbitrarily variable in phase and magnitude 7. The top branch of the T, 9, feeds a crystal detector 10. The opposite horizontal branch of the T is extended into a wave guide 16 which terminates in an open end within a few millimeters of the vibrating surface 13. The vibrations of this surface are in a line parallel to the center line of the wave guide.

In use, the impedance 7 is adjusted so that the variation of the impedance reflected from the open end of the wave guide centers about a point somewhat removed from the null point of the system, being chosen in such a manner that the output of the detector is a linear function of the distance of the wave guide to the reflecting sheet, over as long a range of distance as possible. This linear section of the detector-output-versus-target-distance curve depends on the target geometry, the distance of the wave guide to the target, and the settings of the variable impedance 7. Under optimum settings it will not exceed about 2 millimeters for an output voltage from the crystal of 8 to 10 millivolts. The linear range of this system is small compared to that described in the following, but it has the advantage of being set up in a short time from standard components.

In place of the attenuator 4, a unidirectional transmission line will prevent pulling of the oscillator and yet allow full power to reach the magic T. This results in increased sensitivity and generally better performance. The figure of sensitivity quoted in the foregoing is for this arrangement.

The slope of the reflected-power-versus-target-distance curve varies considerably with the materials of the target but is usable for a wide variety of materials including nonconductors, provided their characteristics are such that a reasonable percentage of the power incident upon them is reflected at the surface.

The system shown in Fig. 3 differs from the preceding one in the nature of the pickup mechanism. The heart of the device, the 10-decibel directional coupler, is not as commonly available as the magic T of the preceding method and, thus, this system is not as easily set up. However, the linear range of this system is about 1/4 inch long, thus considerably exceeding that of the preceding system.

In Fig. 3, a microwave generator 3 (usually a *723 A/B* klystron) couples power into a wave guide, after which it is passed into an attenuator 2 (the writer used about 3 decibels), the purpose of which is to prevent pulling of the oscillator by impedance variations of the load. The unidirectional transmission line referred to previously would probably be an improvement here but, since the 10-decibel coupler is not now conveniently available, it has not been tried. From the attenuator, the power passes down a directional coupler 5, which is set to couple power from reflections. From the directional coupler, the power enters a single screw tuner 9, the function of which is to reflect a portion of the power back to the directional coupler in such a manner as to control the linearity of the device. From the screw tuner, the power passes out through a choke fitting 12, is reflected from the vibrating surface 10, reenters the wave guide, passes back to the directional coupler, combines with the power reflected from the screw tuner, and is detected in the directional coupler output 4.

The action of the system can be traced as follows. Neglect for the moment the screw tuner and the choke fitting on the end of the wave guide facing the vibrating surface. If the end of the wave guide is assumed to be a spherical radiator and the vibrating surface to be a mirror, then, as the mirror recedes from the end of the wave guide, the power reflected back into it will drop off as the square of the distance. As the target further recedes from the end of the wave guide, the power reflected from the target is alternately reinforced by the reflections from the unmatched termination of the wave guide and is cancelled by them. This is shown in Fig. 4, curve *A*.

It is seen that a point of inflection in the curve does not occur until the target is quite far away from the end of the wave guide, thus giving only small sections of the curve which may be considered linear. Now consider a choke flange place on the terminating end of the wave guide. If the choke acts here as it does in the standard line connections, 100-per-cent reflection, or nearly so, should be expected until the target has receded an appreciable distance from the wave guide. This gives the curve *B* in Fig. 4. Here it is seen that a point of inflection occurs in the part of the curve where there is appreciable power being reflected back into the wave guide. However, the linear region is still not as long as desirable. To lengthen the linear region, a screw tuner is inserted in the line to introduce a reflection into the wave guide of such phase as to cancel partially the reflections at the lower part of the curve. The only effect of this reflection on the upper part of the curve, where the reflection is approaching 100 per cent, would be to change the phase. But, as the directional coupler detects only the reflected power without regard to its phase, there

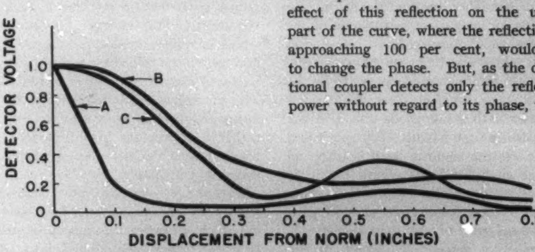


Fig. 4. **Displacement versus output curves, directional and coupler pickup**

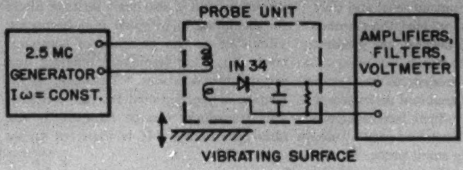


Fig. 5. Mutual inductance field pickup

is very little change in this region. The result is a staisfactorily long linear region, shown in curve *C*, Fig. 4.

The linear range and stability of the directional-coupler system is considerably better than the magic T system previously described. The sensitivity at the detector is approximately 5 millivolts per mil of displacement. To use this device for static calibration against dynamic displacement, either the amplifiers which follow the detector (11, Fig. 2; 6, Fig. 3) must be of the d-c variety or the ouput of the detector must be chopped. The writer, using fairly crude methods, was able to check optical methods of measuring vibrations to 5 per cent or better without difficulty.

Both systems described in this section require a target area at least 1 inch square of flat conducting material. A probe of very small dimensions capable of looking at very small target areas seems possible by going to coaxial lines and tapering them to small dimensions; however, this has not been investigated. In fact, the number of possible probe geometries appears quite large. But it must be remembered that, in almost all cases, the linear range of the probe will always be appreciably less than the probe dimensions. This is true of any electric field pickup, to the best knowledge of the writer.

The sensitivity of a microwave system varies with the configuration used, being greatest in the directional-coupler variety, and is directly proportional to the power fed into the probe by the oscillator. Thus, the use of continuous-wave magnetrons, or sources of greater power than is produced in the *723 A/B* klystron, is promising for this purpose.

The linearity of microwave systems depends on the particular system used, that of the directional coupler type being adjustable within the limits of the linear portion of the curve by the screw tuner. The stability of and freedom from noise and extraneous signals of the system depend almost entirely on the oscillator. The systems are not, however, very compact, and, in spite of the folding of the line and its components, tend toward bulkiness and weight.

Mutual Inductance Systems

National Bureau of Standards System

The first system to be discussed here is that developed by Yates, Davidson and Greenough.[2,3] of the Bureau of Standards. The basic idea is shown in Fig. 5. The mutual inductance of two coils is varied by the presence of a conducting plate, the mutual inductance decreasing as the conducting plate approaches the unit. If now the first coil is supplied with power in such a way that the product of frequency and current is held constant, then the voltage from the second coil depends only on the position of the conducting plate. By suitable arrangement of the two coils and the circuitry supplying them, the target distance versus coil output can be made a linear curve.

At the frequency used here (in the neighborhood of 2 megacycles), the output voltage from the secondary coil would be seriously attenuated by cable capacitance and, therefore, it is desirable to place the detector close to the coil and send only the rectified output back through the probe cable. This is done in the National Bureau of Standards device by the use of a germanium diode mounted in the barrel of the pickup probe.

The geometry and arrangement of the two coils can be varied considerably. However, since it is best to have a target area as large as the largest coil, the size of the coils should be as small as the requirements of the linear region will permit. It should not be expected that the linear region will exceed in length a small part of the coil diameter; a figure of 7 per cent has been suggested. It would seem desirable to have one coil smaller than the other, on a basis of experience, and the separation along their axis may vary from zero to a distance equal to about half the diameter of the larger.

It has been stated that the product of frequency and current must be held constant. To do this, the primary of another pair of coils is inserted for monitoring purposes in series with the lead supplying the primary of the pickup coil with power. The output voltage of the secondary of their pair of coils is rectified, filtered, amplified, and fed back to the screen of the oscillator tube in such a manner as to raise the screen voltage of the oscillator tube, and thus its output, should the product of frequency and current start to decrease.

For system is calibrated statically as follows. A zero position at the start of the linear portion of the curve is selected and the probe is placed in this position. The steady d-c output caused by this position is bucked out by suitable circuitry. The probe is now displaced a known amount, thus giving a net output equal to the change caused by its shift in position. This new d-c output of the probe is chopped, and the resultant square wave is fed to an amplifier, and to a voltmeter set to read peak-to-peak voltages. For further details see reference 3.

Boeing Modification of the National Bureau of Standards System

A frequency-modulated system has been developed which eliminates the effect of cable capacitance between the probe and detector. It is essentially an extensive modification of the National Bureau of Standard mutual inductance pickup and uses the variation of the mutual inductance of the probe to vary the frequency of an oscillator. The oscillator frequency is then fed into a discriminator. The chief advantages of the frequency-modulated system appear to be smaller possible probe sizes, the use of longer cables, and increased stability and sensitivity. At this time, however, the system is still under development and, therefore, there are few experimental data on it. It is being developed as a displacement pickup for reading static displacement; however, there seems to be no apparent difficulty in incorporating the frequency-modulated features into the vibration standard of the National Bureau of Standards.

Summary

Capacitive systems are widespread and varied; they are among the oldest of the methods of picking up vibration. They suffer most from difficulties caused by the capacitance of the leads attached to the sensing element; also there are difficulties inherent in producing a linear system. They are less particular about the conductivity of the material looked at in the vibrating surface, and ferromagnetic materials would cause no trouble with them, but their usefulness on nonconductors seems limited. They require probes of appreciable size. Their output is usually low level.

Microwave systems have as great a possible variety as capacitive systems, and have the advantage that a small geometry change at very high frequencies can effect a large reactance change. They are usable on nonconductors which will give appreciable reflections from their surfaces, and they also will work on magnetic material.

The probe size required at the present time is comparable to other systems, but there seems a possibility that very small probes might be made with a larger linear region, in relationship to probe size, than with the other systems discussed. One variety may be made with standard components available in any radar laboratory. Their chief drawback is bulkiness and low output level and they are probably less stable than other systems, although there is, as present, a lack of data on this point.

Mutual inductance systems are probably the most developed and perfected of any of the methods. They have been engineered for convenience and ease of use, and have moderately small probes on the end of cables to the indicating equipment thus eliminating bulky pickups being placed near the vibrating object. They give high output voltages for small displacements. They cannot be used on nonconducting materials. The National Bureau of Standards system has an established calibrating procedure, permitting direct reading in inches, end to end of the vibration, and it also has a series of filters for examination of small high-frequency vibrations in the presence of comparatively large low-frequency building and apparatus vibrations. The new modification of the National Bureau of Standards transducer can be handled by the same procedures. It is expected to be quite stable.

References

1. Vibration Measuring Device, H. F. Clarke. *U. S. Patent 2,596,531* and *2,596,529*, 1952.

2. Electronic Micrometer, M. L. Greenough. *Electronics*, New York, N. Y., Aug. 1947, pp. 11–13, 26.

3. Wide Range Calibrator for Vibration Pickups, W. Yates, M. Davidson, *Ibid.*, Sept. 1953, pp. 183–85.

Effects of Transverse Compressional Stress on Magnetic Laminations

R. E. FISCHELL
ASSOCIATE MEMBER AIEE

Synopsis: Toroidal, laminated cores of several important magnetic materials were tested for their magnetic properties at compressional stresses ranging from 0 to 500 pounds per square inch (psi). The compressional stress was applied in a direction perpendicular to the plane of the laminations and perpendicular to the direction of the magnetic field. The materials tested were 3.5 and 6.4-per-cent silicon-iron, 81 and 90-per-cent nickel-iron, and both hydrogen- and air-annealed samples of *AEM 4750*, 12 Alfenol, and 16 Alfenol. Tests were run at frequencies of zero, 60, and 400 cycles per second (cps). Under the action of an increasing compressional stress, each of the materials having nonzero magnetostriction exhibited a decrease in maximum permeability and remanence, an increase in coercive force, and a maximum flux density which remained virtually unchanged. The a-c magnetization curve was shifted downward, and the a-c hysteresis loops became less rectangular as the compressional stress was increased. Except at very low fields, materials having zero magnetostriction were unaffected magnetically by the transverse compressional stress. The experimental results indicate that magnetic effects of transverse compressional stress are dependent on the magnitude and independent of the sign of the magnetostriction. The results of this paper can be applied in choosing a magnetic material for a particular application where sensitivity to transverse compressional stress is an important design consideration.

IN MANY important applications, magnetic materials in the form of thin laminations are subjected to a compressional stress. Particularly important examples of this type of application are seen in the clamping of magnetic laminations for electric motors, transformers, and synchro components. In each of these examples, the direction of the compressional stress is perpendicular to the direction of the magnetic field and perpendicular to the plane of the laminations; hence, this particular case warrants special consideration.

The effect of transverse compressional stress on ten different specimens of magnetic materials is described. The ten specimens tested were hydrogen-annealed samples of *AEM 4750* (a nickel-iron alloy), 12 Alfenol, and 16 Alfenol (aluminum-iron alloys), 3.5 and 6.4-per-cent silicon-iron, 81 and 90-per-cent nickel-iron, and air-annealed samples of *AEM 4750*, 12 Alfenol, and 16 Alfenol. None of these materials was grain oriented. The specimens tested were laminated toroids with thin paper insulation between the laminations. The toroids measured $1^5/_8$-inch inside diameter and 2.0-inch outside diameter; the lamination and insulating paper thicknesses measured 0.015 and 0.001 inch respectively.

The effect of stress on magnetic properties when the magnetic field is in the same direction as the applied stress has been described by Bozorth[1] and others. For this case it has been found that the sign of the magnetostriction determines whether permeability increases or decreases with a change in stress. When a compressional stress is applied in the same direction as the magnetic field, the permeability will increase with increasing pressure for a material having negative magnetostriction, and the permeability will decrease if the material has positive magnetostriction.

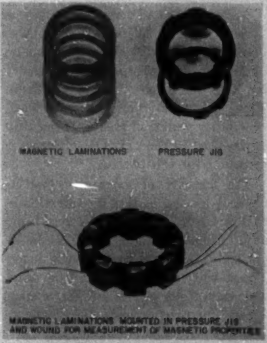

MAGNETIC LAMINATIONS
PRESSURE JIG
MAGNETIC LAMINATIONS MOUNTED IN PRESSURE JIG
AND WOUND FOR MEASUREMENT OF MAGNETIC PROPERTIES

Fig. 1. Special jig for applying a transverse compressional stress to magnetic laminations

---

Paper 55-171, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing December 7, 1954.

R. E. Fischell is with the Naval Ordnance Laboratory, Silver Spring, Md.

Vol. 7
Pt. 1
1955

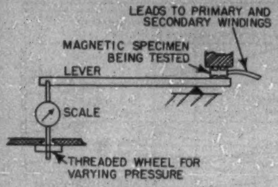

LEADS TO PRIMARY AND SECONDARY WINDINGS
MAGNETIC SPECIMEN BEING TESTED
LEVER
SCALE
THREADED WHEEL FOR VARYING PRESSURE

Fig. 2. Lever and scale system for applying transverse compressional stress to magnetic laminations

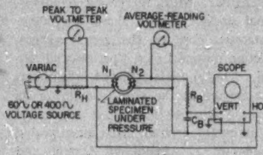

PEAK TO PEAK VOLTMETER
AVERAGE-READING VOLTMETER
VARIAC
N_1
N_2
R_H
60∿ OR 400∿ VOLTAGE SOURCE
LAMINATED SPECIMEN UNDER PRESSURE
R_B
C_B
SCOPE
VERT
HOR

Fig. 3. Circuit for measuring a-c magnetic properties

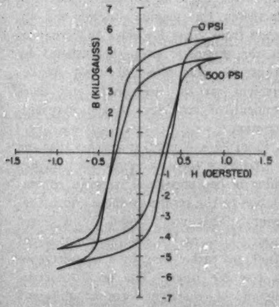


Fig. 4. Dynamic hysteresis loops at a frequency of 400 cps for hydrogen-annealed 16 Alfenol at transverse compressional stresses of 0 and 500 psi

Somewhat different results were observed when the compressional stress was applied perpendicular to the magnetic field. Materials having both negative and positive magnetostriction were tested, and both these classes exhibited a decrease in permeability with increasing pressure. The degree of sensitivity to this transverse compressional stress was dependent on the magnitude of the magnetostriction. It then would be expected that materials having zero magnetostriction should be completely insensitive to pressure changes. Two materials having zero magnetostriction were tested and, except at fields below 0.7 oersted, these materials exhibited essentially no sensitivity to transverse compressional stress. Therefore, the theory that the sensitivity to transverse compressional stress is dependent upon the magnitude of the magnetostriction appears to be substantiated by experiment.

Experimental Procedure

The compressional stress was applied by placing the magnetic laminations in a special fixture, as shown in Fig. 1. By the use of this device, a compressional stress of varying magnitude was exerted on the sample without interfering with the electric windings which are necessary for the determination of magnetic properties.

The device for applying the compressional force was a lever and scale system, as shown in Fig. 2. With this system it was possible to maintain constant pressure on the sample while it was being tested, and it was also possible to increase the compressional stress in increments of 100 psi without relieving the stress while changing from one pressure to another.

Two types of tests were made, and for each test the magnetic properties of the material were measured at compressional stresses of 0, 100, 200, 300, 400, and 500 psi, and the magnetic properties were also measured when the stress was removed. The first of these tests was the standard d-c fluxmeter measurement of maximum permeability, remanence, coercive force, and near saturation flux density, as measured at a field of 30 oersteds.

The second test was made with a-c fields at frequencies of 60 and 400 cps. The purpose of this test was to determine the effect of transverse compressional stress on the a-c magnetization curve and dynamic hysteresis loop. Since the procedure for these a-c measurements is not well known, it will be described briefly here.

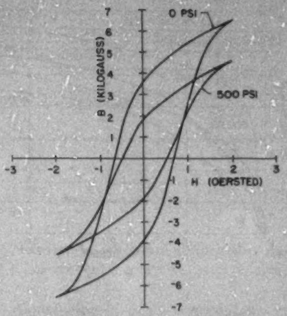


Fig. 5. Dynamic hysteresis loops at a frequency of 60 cps for hydrogen-annealed 90-per-cent nickel-iron at transverse compressional stresses of 0 and 500 psi

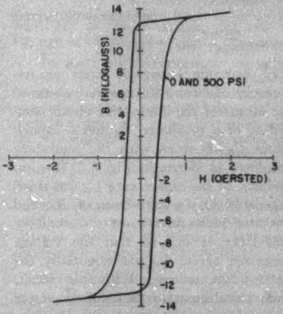


Fig. 6. Dynamic hysteresis loops at a frequency of 60 cps for hydrogen-annealed 6.4-per-cent silicon-iron at transverse compressional stresses of 0 and 500 psi

The circuit for the a-c tests is shown in Fig. 3. The magnetizing field was measured by reading the voltage developed across the metering resistance R_H in the primary circuit. Since the circuit used approximates a voltage source, the current wave is far from being sinusoidal in shape. However, the maximum value of primary current, and therefore H_m, can be determined precisely by means of a peak-reading voltmeter. In terms of the readings of a peak-to-peak voltmeter, H_m is given by

$$H_m = \frac{0.2\pi N_1 V_{p\text{-}p}}{LR_H} \text{ oersted} \qquad (1)$$

where

N_1 = number of primary turns
$V_{p\text{-}p}$ = voltage reading of a peak-to-peak voltmeter
L = mean length of the toroid, centimeters

For toroids of a given mean length L, it is possible and convenient to choose N_1 and R_H so that the peak-to-peak voltmeter reads H_m directly in oersteds.

The value of B_m can be calculated from the voltage reading across the secondary turns N_2 if a rectified average-reading voltmeter is used.[2] This will be true even though (as is actually the case) the flux wave is not sinusoidal in shape. It can be shown that the maximum a-c flux density will be given by

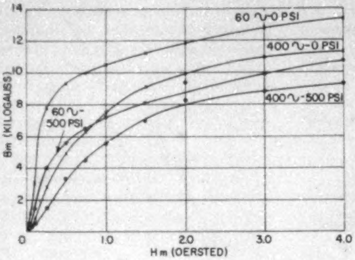


Fig. 7 (left). A-c magnetization curves for hydrogen - annealed AEM 4750

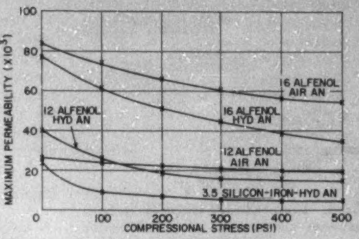


Fig. 9 (right). Maximum permeability (direct current) as a function of transverse compressional stress

$$B_m=\frac{V_{2rms}\times10^8}{\pi\sqrt{2}N_2Af}\text{ gausses}\qquad(2)$$

where

V_{2rms} = voltage as measured by a rectified average-reading voltmeter calibrated in rms voltage of a sine wave

N_2 = secondary turns

A = cross-sectional area of the toroid, square centimeters

f = frequency, cps

The dynamic hysteresis loops at frequencies of 60 and 400 cps were observed by means of the integrating circuit consisting of a resistance R_B and a capacitance C_B in series with the secondary winding, as shown in Fig. 3. The impedance of the capacitance C_B was small compared to the impedance of R_B, and the total series impedance of C_B and R_B was large enough so that the voltage measured across N_2 was essentially the open-circuit voltage. With this circuit, the vertical input on the oscilloscope was proportional to the flux density; and by connecting the horizontal input of the oscilloscope across the metering resistor R_H, the horizontal input was proportional to the magnetic field intensity. Since the values of H_m and B_m were calculated from the voltmeter readings, the vertical and horizontal scales of the oscilloscope were calibrated in terms of kilogausses and oersteds. A calibrated dynamic hysteresis loop was then drawn by photographing the hysteresis loop and using an opaque projector to transfer the photograph onto a sheet of graph paper.

Discussion of Results

Because of limited space, it is not practical to give here all the experimental curves of the various d-c and a-c tests. However, an attempt is made to illustrate with sufficient data the characteristic changes in magnetic properties as a function of transverse compressional stress.

The effect of transverse compressional stress on the a-c properties of a few of the materials tested is shown in Figs. 4 through 8 inclusive. Fig. 4 shows the 400-cps hysteresis loop for the positive magnetostriction material 16 Alfenol at compressional stresses of 0 and 500 psi. The effect of the compressional stress was to cause the dynamic hysteresis loop to become less rectangular. This effect was typical for all the materials tested which had nonzero magnetostriction. Fig. 5 shows the 60-cps hysteresis loop for the negative magnetostriction material of 90-per-cent nickel-iron. When a compressional stress is applied to a material having negative magnetostriction in the same direction as the magnetic field, the exact reverse effect from that shown in Fig. 5 is observed, i.e., the hysteresis loop becomes more rectangular with increasing compressional stress.[3] Fig. 6 shows the 60-cps hysteresis loop for the zero magnetostriction material of 6.4-per-cent silicon-iron. This material showed no measurable change in the size or shape of its dynamic hysteresis loop at values of H_m above 0.7 oersted. The air-annealed samples of 12 Alfenol, 16 Alfenol, and *AEM 4750* showed a smaller change in their a-c hysteresis loops than did the hydrogen-annealed samples of the same material. Since air-annealing in general decreases the magnetostriction, it appears from these a-c data that changes in the magnetic properties caused by transverse compressional stress are dependent on the magnitude and independent of the sign of the magnetostriction.

From Figs. 4 and 5 it appears that the a-c coercive force decreased at a higher value of compressional stress in contradiction to the results of the d-c measurements. This was caused by a lower value of B_m at the compressional stress of 500 psi rather than by a true decrease in coercive force. The accurate comparison of coercive force at different values of compressional stress can be made only when the values of B_m are near saturation and this was not practical to obtain in the a-c tests.

Figs. 7 and 8 show the a-c magnetization curves for hydrogen-annealed *AEM 4750* and 6.4-per-cent silicon-iron. The curve for *AEM 4750* is typical for all

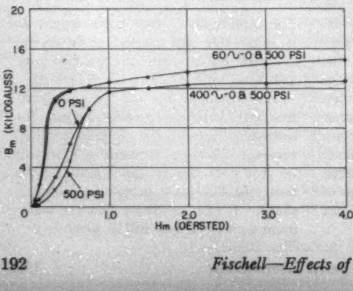


Fig. 8 (left). A-c magnetization curves for hydrogen - annealed 6.4-per-cent silicon-iron

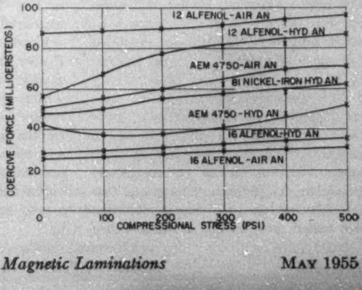


Fig. 10 (right). Coercive force (direct current) as a function of transverse compressional stress

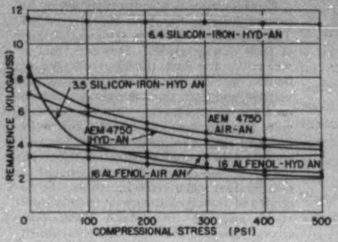


Fig. 11 (left). Remanence (direct current) as a function of transverse compressional stress

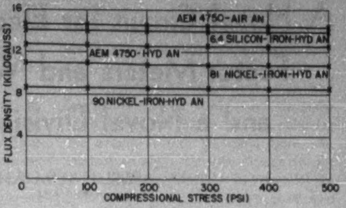


Fig. 12 (right). Flux density at a field strength of 30 oersteds as a function of transverse compressional stress

the materials tested which had nonzero magnetostriction. These curves show that at frequencies of 60 and 400 cps the effect of transverse compressional stress was to decrease the magnetization. Fig. 8, for the zero magnetostriction material 6.4-per-cent silicon-iron shows that, at fields above 0.7 oersted, a change in compressional stress from 0 to 500 psi produced no measurable change in the magnetization curve. A behavior very close to this was observed for the zero magnetostriction material 81-per-cent nickel-iron. Since the magnitude of the magnetostrictive strain varies with field intensity, it is possible that some of the variation in the magnetization curves of 6.4-per-cent silicon-iron and 81-per-cent nickel-iron at fields below 0.7 oersted is caused by the material acquiring some magnetostriction properties at these low field intensities.

The variation of d-c magnetic properties with transverse compressional stress for several of the materials tested is shown in Figs. 9 through 12 inclusive. These figures show that, under the action of an increasing transverse compressional stress, the maximum permeability decreased, the coercive force increased, the remanence decreased, and the near saturation flux density, measured at a field of 30 oersteds, remained virtually unchanged. In the measurements of coercive force and remanence, a maximum field strength of 30 oersteds was used.

Since maximum permeability and coercive force occur at very low field intensities for 6.4-per-cent silicon-iron and 81-per-cent nickel-iron, the changes in these d-c magnetic properties with compressional stress is in line with the changes observed in the a-c properties of these materials at low magnetic field intensities. With the exception of *AEM 4750*, the air-annealed samples showed a smaller percentage change in d-c magnetic properties than did the hydrogen-annealed samples of the same material. The negative magnetostriction material 90-per-cent nickel-iron changed in d-c magnetic properties in the same direction as did the material which had positive magnetostriction. The results of the various a-c and d-c tests are summarized in Table I.

After the stress was removed, the magnetic properties of most of the materials tested returned within a few per cent to their original values. A notable exception was in the measurement of maximum permeability and coercive force for 81 and 90-per-cent nickel-iron. After the stress was removed, the maximum permeability of the 81-per-cent nickel-iron showed a decrease of 9.7 per cent and the coercive force increased 6.5 per cent; the 90 per cent nickel-iron showed a 35-per-cent decrease in maximum permeability and a 23-per-cent increase in coercive force.

Table I. Percentage Change in Magnetic Properties Resulting from an Increase in Transverse Compressional Stress from 0 to 500 Psi

			Percentage Change from 0 to 500 Psi					
Material	Anneal	Magnetostriction Constant	B_m (60 Cycles)	B_m (400 Cycles)	B_m (d-c)	Maximum Permeability	Remanence	Coercive Force
12-Alfenol	hydrogen	+	38.3	35.2	6.42	64.2	28.1	48.4
12-Alfenol	air	+	21.0	20.3	5.28	27.1	18.5	10.0
16-Alfenol	hydrogen	+	18.4	17.8	5.13	58.6	30.9	27.1
16-Alfenol	air	+	7.42	7.20	1.46	34.9	14.0	26.7
AEM-4,750	hydrogen	+	27.7	20.1	2.17	79.4	47.5	20.9
AEM-4,750	air	+	24.1	17.6	2.87	67.0	49.7	41.5
3.5-per-cent silicon-iron	hydrogen	+	40.4	20.8	0.27	81.9	75.4	31.2
6.4-per-cent silicon-iron	hydrogen	0	0	0	0.86	9.25	2.48	12.5
81-per-cent nickel-iron	hydrogen	0	1.04	1.54	2.73	51.4	24.9	29.4
90-per-cent nickel-iron	hydrogen	−	25.7	16.6	3.00	50.8	65.0	4.88

Conclusion

Several conclusions can be drawn regarding the sensitivity of magnetic materials to a transverse compressional stress.

1. Variations in magnetic properties are dependent on the magnitude and not on the sign of the magnetostriction constant.
2. For materials having nonzero magnetostriction, under the action of an increasing transverse compressional stress, the maximum permeability decreased, the coercive force increased, the remanence decreased, the flux density near saturation remained virtually unchanged, and, for a constant value of H_m, the maximum a-c flux density decreased.
3. The magnetic characteristics of materials having zero magnetostriction remain virtually unaltered by transverse compressional stress except at very low fields.

The results of this paper can be applied in choosing a magnetic material for a particular application where sensitivity to transverse compressional stress is an important design consideration. In the manufacture of precision components using magnetic laminations, differences in clamping pressures may cause variations in magnetic properties if the material used has a high sensitivity to transverse compressional stress. Variations in magnetic properties from one set of laminations to another can be minimized by using stress-insensitive magnetic materials. The high values of maximum permeability and saturation flux density of 6.4-per-cent silicon-iron make it a particularly interesting magnetic material where stress insensitivity is desired.

References

1. FERROMAGNETISM (book), R. M. Bozorth. D. Van Nostrand Company, New York, N. Y., 1951, pp. 595–610.

2. A FLUX VOLTMETER FOR MAGNETIC TESTS, G. Camilli. *AIEE Transactions*, vol. 45, 1926, pp. 989–95.

3. Bozorth, *op. cit.*, p. 596.

l. 74
1
55

A Novel Circuit for Electronic Power-Factor Meters and Wattmeters, and a Novel Division Device

M. ABDEL-HALIM AHMED
MEMBER AIEE

Synopsis: Three novel electronic measuring instruments are presented, which may be used to measure power, power factor (pf), voltage, current, or impedance. They use two tubes, operated as cathode followers, in a push-pull fashion, so that the inherent property of a wide-range linear characteristic is employed to advantage. The new principle underlying the operation of the tubes involves the commutation of their currents in alternate half-cycles by means of a square-wave generator. The resulting average plate current is a direct measure of the product $I \cos \varphi$. A rectifier gives a direct current proportional to the alternating voltage, and an electrodynamic instrument multiplies the two direct currents, thus giving a direct reading of the power. For the direct reading of pf, two direct currents are obtained, one being proportional to the product $V \cos \varphi$ and the other proportional to V. Their quotient is a direct measure of $\cos \varphi$. The division is effected by means of a twin logarithmic d-c instrument having a movable scale. The instruments are characterized by their simplicity, reliability, high accuracy, and stability. They have exceptionally good features and predictable performance over a frequency range extending from the lowest power frequencies to a predicted value of about 30 megacycles (mc) when using high-frequency tubes; the main limiting factor being the stray capacities across the output of the square-wave generator. There are, however, some other factors which tend to lower this operating limit, and 10 mc is considered a reasonable higher limit.

THE DIRECT measurement of power at power frequencies and the lower audio frequencies is most conveniently carried out by using an electrodynamic instrument. Alternatively, a double-magnet induction instrument may be used at power frequencies. For the measurement of small amounts of power, or for audio- and high-frequency work, these instruments are not suitable, owing to their well-known inherent drawbacks. For such cases, many indirect methods have been devised, but even then, the pf of the circuit has to be already known.

Many investigators have suggested different instruments which they have claimed to be suitable for the direct measurement of power at high frequencies. A number of the early attempts involve the use of two similar square-law elements, which may be thermal (hot-wire or thermocouple elements) or electronic tubes working in the square-law range of their characteristics.[1–4] In a later attempt[5] the tubes were operated in their logarithmic range. All these devices have the following disadvantages: the two nonlinear elements have to be exactly matched, which is difficult to perform; replacement of the elements will affect the calibration; and the circuits require frequent adjustment. As regards the electronic tubes, even when matched, the desired relation of their characteristics holds only over a limited range. Voltage dividers will thus be unavoidable, and these cause undue insertion loss in the circuit, and trouble at high frequencies. Moreover, the dissipation loss of the current circuit has to be high if any reasonable accuracy is to be expected, especially with the thermal devices. Supply voltage variations, besides causing calibration errors, require frequent adjustment of the zero setting. For all the square-law devices, the insertion unit is of T- or π-configuration, with which it is difficult to provide a satisfactory ground system.

Among others, there are the following three distinct classes of device:

1. *Devices using electrostatic methods.* These are not handy and are too delicate. Difficulty is encountered in their adjustment and setting, and they require high voltages for their operation.

2. *Devices using oscillographs.* These either measure instantaneous values of power or give an indirect measure of the average value through tedious and inaccurate operations. They also suffer from lack of sensitivity.

3. *Devices using multielectrode tubes.*[6–8] The tubes are arranged to operate on the linear parts of their various grid characteristics, which are too small to avoid rectification, and the instruments were found to give readings with current or voltage alone, even with push-pull operation. Discrepancies in the characteristics of individual tubes of a given type are serious, and the choice of matched tubes is too difficult. Frequent check calibration and zero-setting adjustment are necessary. The use of a voltage divider, with its inherent disadvantages, is unavoidable.

To this class a new instrument having improved features has been recently added.[9] The tubes operate in the exponential range of their characteristics. However, this device still suffers from the disadvantages of a limited operating range for the control grid voltage, and from the fact that it cannot be used for circuits having voltages below 20 volts (transformers can be used at low frequencies since the input is of high impedance, but this is not a possible solution at the higher frequencies). The requirement of recalibration when replacing tubes, the necessity of stabilized-supply or battery operation, and the unavoidable rectification errors (though claimed to be small) resulting from the discrepancy of the parameters of the tubes themselves from the requisite values under this mode of operation, are all inherent disadvantages.

There is no known method for the direct measurement of the power factor, $\cos \varphi$, at frequencies above power frequencies. However, the phase displacement φ may be measured by the commonly used phase detectors, in which two constant-amplitude square-wave signals are applied to two elements of a tube. The output produced is a function of the phase difference between the square-wave signals. If the instrument is to be calibrated to read pf the scale would be crowded in its lower part. Besides, this type of instrument is essentially complicated.

The instruments described in this paper use two tubes operating in push-pull and working mainly on the linear parts of their characteristics, which usually cover a wide range. This range can be extended and made as linear as desired by suitable adjustments of the external circuit, thus rectification errors are almost eliminated and voltage dividers are dispensed with. Critical adjustment of the circuit elements is not a necessity in this mode of operation. Negative current feedback is used with advantage, which results in extended linear range, increased stability, less trouble with tube replacements, and insensitivity of calibration and zero setting to supply voltage variations.

The instruments are characterized by their simplicity in theory, circuit design, operation, and adjustment. Almost any tube type can be used, and the performance can be readily predicted from knowledge of the normal tube parameters and circuit elements. A unique main feature lies in the fact that individual d-c components which are proportional to the individual a-c compo-

---

Paper 55-1, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted January 28, 1954; made available for printing October 27, 1954.

M. Abdel-Halim Ahmed is with Cairo University, Giza, Cairo, Egypt.

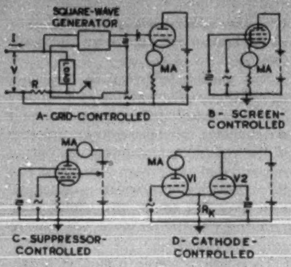


Fig. 1. Principle of the I cos φ meter

nents are available in the circuit, so that direct readings of voltage, current, pf, and power can be readily obtained with the same instrument with complete reliability and stability. Furthermore, the instruments are suitable for measurements at very low pf. They will therefore be found useful in measuring impedance/frequency and pf/frequency characteristics of a network or apparatus under actual operating conditions.

Among the good features of the instruments are the uniform scale of the indicating instrument, the low insertion loss of the series resistor in the current circuit, the high accuracy, and the simplicity of adapting the instrument for multirange use.

For unavoidable reasons relating to the lack of certain equipment, stress was given to obtaining the maximum possible output power from the tubes, which forced the use of normal receiving tubes. However, with the use of high-frequency tubes, it is shown that a frequency range comparable with any other type of instrument can be covered. Predicted higher limits of about 30 mc at unity pf and 3.5 mc at zero pf were obtained. Nevertheless, factors not considered are expected to cause trouble at the higher frequencies, especially with the square-wave generator, and 10 mc is considered a reasonable higher limit. The errors resulting from other possible sources are discussed and found remarkably small.

Three types of the instruments have been developed: the grid-controlled type; the suppressor-controlled type; and the cathode-controlled type.

General Outline

The main idea of these commutating type of instruments lies in using an electronic tube having a fairly linear transfer characteristic of control-grid voltage versus plate current, which is readily obtained through the use of negative current feedback. Then a signal is applied to the control grid, and the tube current is commutated only during a proper interval of its cycle. The commutation period is controlled by applying a square-wave signal having the same frequency as the control-grid signal, and of sufficient amplitude, to one of the tube grids, thus allowing the plate current to flow during only one half of the square-wave cycle. The average value of the plate current follows, linearly, the signal amplitude and the cosine of the phase angle between the signal and the square wave. The square-wave voltage is obtained from a square-wave generator designed so that its output is of the same frequency as, and in phase with, the voltage applied to its input.

Fig. 1(A) shows the circuit arrangement for the measurement of pf, using a triode. The voltage V across the load is applied to the input of the square-wave generator, the square-wave output of which is applied to the control grid of the tube. The voltage drop IR across a small resistance R in which the load current passes, is applied to the control grid of the tube, superposed on the square wave. The change in the average value of the output current will be proportional to $I \cos \varphi$, φ being the phase angle between V and I. If I is known or measured (which can be done by the same instrument), or if the voltage IR is adjusted to a predetermined value, of if a suitable dividing circuit is added to give the quotient of the two quantities $I \cos \varphi$ and I, the device will give a direct reading of $\cos \varphi$.

It is obvious that if the input to the square-wave generator were derived from the load current circuit, $\cos \varphi$ would be unity, and the device an ammeter with a linear scale. In the same manner, if the signal applied to the grid of the tube were derived from the voltage circuit, the device would be a voltmeter with a linear scale.

Now, for the measurement of power, the product $I \cos \varphi$ has already been obtained in the form of a direct current, and it now remains to obtain a direct current proportional to V through the use of a conventional rectifying circuit. The two direct currents are multiplied most easily by the use of an ordinary electrodynamic instrument, and the device is then a direct-reading wattmeter with a linear scale.

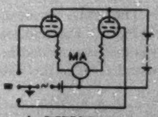
That it is not possible to connect the tube d-c supply voltage to ground is the main inconvenience noticed in the grid-controlled circuit. However, in many instances this is not a serious defect, especially if it is noted that this voltage is raised above ground by only a fraction of a volt, the usual value of the voltage drop IR.

In some circuit arrangements, especially in high-frequency work, and in order to provide a suitable grounding point for the circuit, it may be found more convenient to apply the two signals to different grids of the tube. In Fig. 1(B) the square-wave signal is applied to the screen grid of a pentode whose suppressor grid is internally connected to the cathode. With this procedure, severe requirements are imposed on the square-wave generator, which must provide a positive wave of large magnitude and be capable of supplying the screen-grid current of the tube; moreover, if a push-pull arrangement is used, both sides of its output should give exactly equal amplitudes of the square wave.

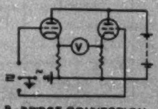
In Fig. 1(C) the square wave is applied to the suppressor grid of a pentode. It requires a square-wave generator giving a negative signal, and thus precludes all these difficulties.

In Fig. 1(D) two tubes are used having a common cathode resistor, which may be triodes or multigrid tubes. Each signal is applied to the control grid of one tube. When $V2$ is conducting, the voltage drop across the common cathode resistor R_K will cause $V1$ to be biased beyond cut-off; when $V2$ is cut-off, $V1$ will be conducting. This circuit has the advantage that it requires a square wave of small amplitude, and will consequently be suitable for use at much higher frequencies.

All circuit arrangements are well adapted for use in a push-pull fashion. This is shown in Fig. 2 for the grid-controlled type. A differential-wound d-c indicating instrument can be used as in Fig. 2(A), or a normal instrument may be used as in Fig. 2(B). It is obvious that in the differential connection the instrument should have a low impedance to the al-



A- DIFFERENTIAL CONNECTION



B- BRIDGE CONNECTION

Fig. 2. Principle of the push-pull I cos φ meter

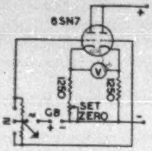


Fig. 3. The grid-controlled I cos φ meter

ternating current. Its coils may be conveniently shunted with capacitors, provided their d-c resistances are relatively low. For the bridge connection, the circuit between the two cathodes should offer high impedance to both alternating and direct currents. When the resistance of the instrument is low, an external resistance may be added in series with it, and the coil may also be shunted with a capacitor.

The use of a push-pull circuit has the merits of obviating the necessity of compensating the zero-signal tube current, reducing zero-setting variations with supply voltage variations, reducing the errors resulting from any possible nonlinearity in the tube characteristics and, finally, increasing the sensitivity of the instrument, thus reducing the insertion loss of the series resistance in the load current circuit.

The merits of the cathode-follower circuit are well known. Its main advantages relating to the present type of work lie in providing a wide-range linear-transfer characteristic, whose slope depends mainly on the value of the cathode resistor and is almost independent of the tube parameters, reducing tube current variations with supply voltage variations. Therefore no readjustment or recalibration is necessary when replacing tubes, no readjustment of zero setting is necessary during operation, and the stability of the readings is highly increased.

Grid-Controlled Commutating PF Meter

CIRCUIT

The actual circuit of this type is shown in Fig. 3. For zero setting, a variable resistance is provided in one of the cathode circuits. The d-c characteristic for one triode section is shown in Fig. 4. It has a linear range beginning from −4 volts upward. The upper limit of the useful working range is about +20 volts, at which point the grid-cathode potential is zero. The operating point is chosen at about the middle of the useful working range, hence a positive bias (i.e., connected with its positive side toward the grid of the tube) of about 8 volts is required. An a-c signal up to a 12-volt peak (8.5 volts rms) can be applied to the control grid.

The push-pull square-wave generator has to carry the grid potential in the negative direction so that the tube is cut off with the maximum value of the a-c signal, and thus the square-wave generator will give, from each side of its output, zero potential during 1/2 cycle and a negative potential of more than 32 volts during the other 1/2 cycle. It is not essential that both sides should give equal negative voltages.

SQUARE-WAVE GENERATOR

The circuit of the square-wave generator, which is simple and direct, is shown in Fig. 5. Clipping is done by the double-diode clipper tubes as well as by the amplifier stages. The amplifier stages work with almost zero grid bias at no signal, and when the signal is applied, they are automatically biased to about the peak value of the signal through grid-leak rectification action. The negative peak value of the signal is sufficient to bring the amplifier tube to cutoff. By this means, the amplifier stages help also in clipping off the higher portions of the signal, in which the effect of stray capacities appears most, leaving the signal with steep sides. Thus by these successive clipping operations, the effects of stray capacities can be neglected in all the stages prior to the grids of the output tubes. Effects of stray capacities in the output circuit, however, will remain as the main limiting factor of the frequency range in which the square-wave generator may be used, since no clipping stage follows.

That the signal applied to the grid of a tube must be more than sufficient to bring it to cutoff is most important in the output stage, since it is extremely necessary for the satisfactory operation of the pf meter that both tubes will begin from the same predetermined working point, and thus the square-wave generator has to give a signal which falls down to exactly zero.

It is possible to extend the frequency range through the addition of a clipping stage, the use of high-frequency tubes for the pf meter, minimizing stray capacities of the connections, reducing the output resistance of the square-wave generator through the use of high plate voltages and more powerful tubes, and adding compensating inductances in the square-wave generator circuit.

There are, however, a number of factors which cause trouble at high frequencies in the square-wave generator, and therefore tend to lower its useful upper frequency limit further than would be predicted by consideration of the effect of stray capacities across the output circuit alone. Such factors are, for example, the inductance in tube element leads, the cathode-to-ground capacitance, balance on clipping diodes, and phase shift in clipping stages. Balance on the clipping diodes is, however, not an essential requirement, since it is not necessary that the push-pull output be of the same magnitude on both sides, as long as the magnitudes are enough to bring the commutating tubes to cutoff with the maximum value of applied a-c signal voltage.

Trouble from phase shift in the various stages of the square-wave generator can be eliminated by adding at its input a compensating simple resistance-capacitance phase-shifting circuit. Before measurements at high frequencies are started, the instrument is switched so as to read either *V* or *I*, and the compensating circuit is adjusted for maximum reading on the indicating instrument.

The square-wave generator is suitable for use with input voltages covering a very wide range, from a fraction of a volt to several hundred volts, and can therefore be connected directly to the load voltage circuit without the necessity of using a voltage divider. The input circuit has a high impedance which causes no appreciable insertion loss. However, if a much higher input impedance is desired, or if it is required to work with very small input voltages, a single-tube conventional amplifier stage can be added at the input. This may use a cathode-follower circuit if the input voltage is not too low.

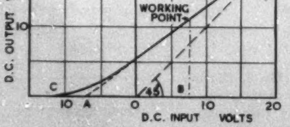
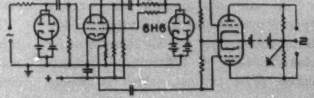


Fig. 4 (left). Static characteristic for one triode section of Fig. 3

Fig. 5 (right). Square-wave generator for use with the grid-controlled I cos φ meter



The square-wave generator must give equal intervals of its two output waves. This can be easily checked by applying the square-wave voltages alone to the pf meter, which should then show no indication. Errors in timing may be caused by harmonics existing in the input voltage when these harmonics are not in phase or in antiphase with the fundamental. These errors increase with the harmonic per cent (%) and the phase difference between the harmonic and the fundamental, and decrease for harmonics of higher order.

To have an idea of the error in timing produced by harmonics that occur in practical cases, consider, for example, the case of a class-*C* amplifier driven from a sinusoidal source and the output is a tuned load. This circuit operates with almost unity pf. It can be shown that the ratio of the tuned load impedance Z_h at the frequency of the *n*th harmonic, to its impedance Z_f at the fundamental frequency, is given by

$$\frac{Z_h}{Z_f}=\frac{1}{1+jQ\dfrac{n^2-1}{n}}$$

where Q represents the Q-factor of the tuned load. With a pessimistic low value of Q, say 20, the foregoing ratio would amount to 3.33% for the second harmonic, 1.87% for the third, and so on, while the phase angle is almost 90 degrees. With a current flow of 100 degrees in the class-*C* amplifier and three-halves law, the current flowing in the tuned load will have 83.5% second harmonic and 57% third harmonic. Therefore the voltage developed across the tuned load will have harmonics of the order of 2.78% and 1.065% respectively. These will produce an error in timing of about 1.6 degrees and 0.6 degree respectively.

It should still be pointed out that a timing error or a phase shift which arises from an even-order harmonic can always be totally eliminated by applying a small adjustable direct voltage at the input of the square-wave generator, superposed on the alternating voltage.

OPERATION OF PF METER

Ideal Operation

In this case it is assumed that the two tubes have identical and exactly linear transfer characteristics down to the cutoff point. The magnitudes of the applied signals are such that in the half-cycle when the square-wave voltage is zero, the a-c signal will drive the grid of the tube neither more positive than the cathode nor more negative than the cutoff point; and in the other half-cycle, when the square-wave voltage is negative, the a-c signal will not drive the grid more positive than the cutoff point.

Let S denote the slope of the transfer characteristic of either tube which can be as

$$\frac{\mu}{\mu+1}\,\frac{1}{R_k+\dfrac{\rho+R_a}{\mu+1}}$$

where

μ = amplification factor
ρ = plate internal impedance
R_a = plate load impedance
I_0 = steady no-signal plate current of either tube
T = time period of 1 cycle $=1/f=2\pi/\omega$, f being the frequency
$F\sin(\omega t-\varphi)$ = a-c signal applied to the grid, φ being the phase angle between this signal and the square-wave signal
I_1 and I_2 = average currents of the two tubes when the signals are applied

then

$$I_1=\frac{1}{T}\int_0^{T/2}[I_0+SF\sin(\omega t-\varphi)]dt$$
$$=\frac{I_0}{2}+\frac{S}{\pi}F\cos\varphi$$

$$I_2=\frac{1}{T}\int_{T/2}^{T}[I_0+SF\sin(\omega t-\varphi)]dt$$
$$=\frac{I_0}{2}-\frac{S}{\pi}F\cos\varphi$$

and the average current difference is given by

$$I=I_1-I_2=(2S/\pi)F\cos\varphi \qquad (1)$$

which is proportional to the product of the magnitude of the applied a-c signal and the pf.

Effect of Different Transfer Characteristic Slopes

In a cathode-follower circuit, the tube current is almost independent of the tube parameters, and the slope of the transfer characteristic S is mainly defined by the value of R_k. However, if for any reason the two tubes give different transfer characteristic slopes, say S_1 and S_2, the reading of the indicating instrument will be

$$I=\frac{S_1+S_2}{\pi}F\cos\varphi \qquad (2)$$

i.e., the average value $(S_1+S_2)/2$ is to be considered instead of S, and the reading is still proportional to the product of the magnitude of the a-c signal and the pf.

Effect of Error in Square-Wave Timing

It has been stated that the square-wave generator can, in most cases, be adjusted to give exact timing. However, it is interesting to show that even without such an adjustment, the error produced in the pf meter indication is very small.

If T_1 and T_2 are the time intervals of the square wave, so that $T_1+T_2=T$, and writing $\omega T_1=\pi+\delta$, $\omega T_2=\pi-\delta$, or $\omega(T_1-T_2)=2\delta$, then, when the square-wave signal only is applied, the indicating instrument reads $I_0(T_1-T_2)/T=I_0(\delta/\pi)$.

When the signal $F\sin(\omega t-\varphi)$ is applied

$$I=\frac{1}{T}\int_0^{T_1}[I_0+SF\sin(\omega t-\varphi)]dt-$$
$$\frac{1}{T}\int_{T_1}^{T}[I_0+SF\sin(\omega t-\varphi)]dt$$
$$=I_0\frac{\delta}{\pi}+\frac{S}{\pi}F[\cos\varphi-\cos(\omega T_1-\varphi)]$$

and the change in the reading is then given by

$$\frac{S}{\pi}F[\cos\varphi-\cos(\omega T_1-\varphi)]$$
$$=\frac{S}{\pi}F[\cos\varphi+\cos(\varphi-\delta)] \qquad (3)$$

Compare this with equation 1, and the error in the reading will be

$$\frac{S}{\pi}F[\cos\varphi-\cos(\varphi-\delta)]$$

and the relative error is

$$\frac{\cos\varphi-\cos(\varphi-\delta)}{2\cos\varphi}=\frac{1}{2}\tan\varphi\sin\delta-\sin^2\frac{\delta}{2}$$

and, since δ is always very small

$$\text{relative error}\cong\frac{\delta}{2}\tan\varphi-\frac{1}{4}\delta^2$$

This error would naturally be infinite at zero pf, since the true reading should have been zero. However, this is no indication of the error, and the error in indicating instruments is usually, and preferably, judged by referring to the maximum instrument reading, $(2S/\pi)\,F_{\max}$; hence

full-scale relative error

$$=\frac{F}{F_{\max}}\,\frac{\cos\varphi-\cos(\varphi-\delta)}{2}$$

and for the extreme condition when $F=F_{\max}$ (=12 volts)

full-scale relative error

$$=\frac{1}{2}[\cos\varphi-\cos(\varphi-\delta)]$$
$$=\sin\frac{\delta}{2}\sin\left(\varphi-\frac{\delta}{2}\right)\cong\frac{\delta}{2}\sin\left(\varphi-\frac{\delta}{2}\right)$$

For the previously estimated value of $\delta=1.6$ degrees, this error amounts to 0.02% at unity pf and 1.38% at zero pf.

Further, consider a much exaggerated value of 10% third harmonic, the timing error in this case is 5.5 degrees, which, in the extreme case when $F=F_{\max}=12$ volts,

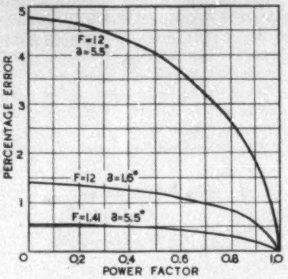


Fig. 6. Full-scale relative error attributable to imperfect square-wave timing

produces a relative error of only 0.23% at unity pf and 4.8% at zero pf. Yet, for a small insertion loss of the current element in the load circuit, the a-c signal should never exceed 1 volt rms, i.e., $F=1.41$ volts, and in this case the relative error amounts to 0.027% and 0.565% respectively.

Fig. 6 shows this error for different pf in the three cases considered in the foregoing, from which it is concluded that this error, if it exists, is always small or negligible.

Effect of Nonlinearity in Transfer Characteristics

Consider first that the nonlinearity occurs all over the working range of the characteristic, then this latter can be expressed in the form

$$i=I_0+Sv+Pv^2$$

This expression assumes symmetrical deviations from a straight-line relation, since it neglects the cubed-voltage term and terms of higher order. This assumption is justified mainly owing to the large amount of negative feedback present, and also owing to the relatively low value of a-c signal voltage that would be allowed compared to the linear part of the characteristic (Fig. 4). It is well known that negative feedback considerably reduces the nonlinear terms of the characteristic, even when large excursions on the characteristic are considered, and that the amount of reduction increases rapidly with the order of the nonlinear term.

When an a-c signal of the form $v=F\sin(\omega t-\varphi)$ is applied, then

$$I_1=\frac{1}{T}\int_0^{T/2}[I_0+SF\sin(\omega t-\varphi)+PF^2\sin^2(\omega t-\varphi)]dt$$

$$=\frac{I_0}{2}+\frac{S}{\pi}F\cos\varphi+\frac{PF^2}{4}+\frac{PF^2}{4\pi}\sin 2\varphi$$

$$I_2=\frac{1}{T}\int_{T/2}^{T}[I_0+SF\sin(\omega t-\varphi)+PF^2\sin^2(\omega t-\varphi)]dt$$

$$=\frac{I_0}{2}-\frac{S}{\pi}F\cos\varphi+\frac{PF^2}{4}+\frac{PF^2}{4\pi}\sin 2\varphi$$

and the instrument reading will be

$$I=\frac{2S}{\pi}F\cos\varphi \tag{4}$$

This is the same as equation 1, and thus no error is produced, a result to be expected from the push-pull operation of the circuit.

It is more common for the transfer characteristic to be linear along its upper part, and to show nonlinearity only in the lower part. In an exaggerated case where the nonlinearity is assumed to extend up to the static working point of the characteristic (when this is not the case, a smaller error would be expected), this latter can be divided into two portions, which may be expressed as $i=I_0+Sv$ for positive values of the input voltage v, and $i=I_0+Sv+Pv^2$ for the negative values of v.

Now, for the first tube, which is conducting in the time period 0 to π/ω, the applied voltage is $v=F\sin(\omega t-\varphi)$, which will then be negative in the time period 0 to φ/ω and positive in the time period φ/ω to π/ω. The average current of this tube is thus given by

$$I_1=\frac{1}{T}\int_{\varphi/\omega}^{\pi/\omega}[I_0+SF\sin(\omega t-\varphi)]dt+\frac{1}{T}\int_0^{\varphi/\omega}[I_0+SF\sin(\omega t-\varphi)+PF^2\sin^2(\omega t-\varphi)]dt$$

$$=\frac{I_0}{2}+\frac{S}{\pi}F\cos\varphi+\frac{PF^2}{4\pi}\varphi+\frac{\sin 2\varphi}{2}$$

For the second tube, which is conducting in the period π/ω to $2\pi/\omega$, the voltage $v=F\sin(\omega t-\varphi)$ is positive in the time period π/ω to $(\pi/\omega+\varphi/\omega)$, and negative in the time period $(\pi/\omega+\varphi/\omega)$ to $2\pi/\omega$, hence

$$I_2=\frac{1}{T}\int_{\pi/\omega}^{\frac{\pi}{\omega}+\frac{\varphi}{\omega}}[I_0+SF\sin(\omega t-\varphi)]dt+\frac{1}{T}\int_{\frac{\pi}{\omega}+\frac{\varphi}{\omega}}^{2\pi/\omega}[I_0+SF\sin(\omega t-\varphi)+PF^2\sin^2\omega t-\varphi)]dt$$

$$=\frac{I_0}{2}-\frac{S}{\pi}F\cos\varphi+\frac{PF^2}{4\pi}\left(\pi-\varphi+\frac{\sin 2\varphi}{2}\right)$$

The instrument indication will then be

$$I=\frac{2S}{\pi}F\cos\varphi\left[1+\frac{P}{4S\cos\varphi}F\left(\varphi-\frac{\pi}{2}\right)\right] \tag{5}$$

Hence the relative error resulting from nonlinearity is given by

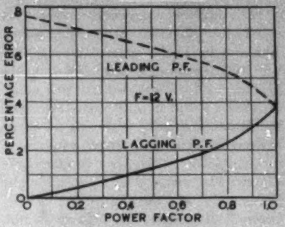


Fig. 7. Full-scale relative error attributable to characteristic nonlinearity

$$\text{relative error}=-\frac{PF}{4S\cos\varphi}\left(\varphi-\frac{\pi}{2}\right)$$

This is seen to be proportional to the amplitude of the applied signal, and is not infinite at zero pf.

The ratio P/S can be shown geometrically to be equal to $\overline{CA}/\overline{CB}^2$ (see Fig. 4), which amounts to 0.008. If an extreme case is considered, when $F=12$ volts, the relative error is shown in Fig. 7 for different power factors. It should be pointed out that although these values are small, they are still too much exaggerated in two ways, the first in considering the nonlinearity to extend up to the steady no-signal working point, and the second in taking $F=12$ volts, while it does not usually exceed a fraction of a volt.

Effect of Harmonics in Load Current Wave Form

Let the a-c signal voltage, in this case, be expressed in the form

$$v=F_1\sin(\omega t-\varphi)+F_n\sin(n\omega t-\varphi_n)$$

where n is the order of the harmonic.

Consider a perfect square wave and a linear characteristic; then

$$I_1=\frac{1}{T}\int_0^{\pi/\omega}[I_0+SF_1\sin(\omega t-\varphi)+SF_n\sin(n\omega t-\varphi_n)]dt$$

$$=\frac{I_0}{2}+\frac{S}{\pi}F_1\cos\varphi \text{ when } n \text{ is even, or}$$

$$\frac{I_0}{2}+\frac{S}{\pi}F_1\cos\varphi+\frac{S}{\pi}\frac{F_n}{n}\cos\varphi_n \quad \text{when } n \text{ is odd}$$

and

$$I_2=\frac{1}{T}\int_{\pi/\omega}^{2\pi/\omega}[I_0+SF_1\sin(\omega t-\varphi)+SF_n\sin(n\omega t-\varphi_n)]dt$$

$$=\frac{I_0}{2}-\frac{S}{\pi}F_1\cos\varphi \text{ when } n \text{ is even, or}$$

$$\frac{I_0}{2}-\frac{S}{\pi}F_1\cos\varphi-\frac{S}{\pi}\frac{F_n}{n}\cos\varphi_n \quad \text{when } n \text{ is odd}$$

and the instrument indication will thus be

$$I=\frac{2S}{\pi}F_1\cos\varphi \text{ when } n \text{ is even, or}$$

$$\frac{2S}{\pi}F_1\cos\varphi+\frac{2S}{\pi}\frac{F_n}{n}\cos\varphi_n \text{ when } n \text{ is odd}$$

It is thus seen that the even harmonics produce no effect whatever. The odd harmonics, however, produce an error whose relative value is given by

$$\frac{F_n}{nF_1}\frac{\cos\varphi_n}{\cos\varphi}$$

For a given value of cos φ, this error is maximum when $\varphi_n=0$, and the full-scale relative error (referred to the maximum instrument reading, i.e., at unity pf) will then be F_n/nF_1. Hence, for an exaggerated value of 10% harmonic content, the maximum error amounts to 3.33% for the third harmonic, 2% for the fifth harmonic, etc.

It is therefore concluded that errors produced by harmonics, whether in the current or voltage wave forms, are either zero or negligible.

Effect of Stray Capacities

The square-wave generator forms that part of the circuit which limits the useful frequency range. The presence of stray capacities across its output represents the main factor limiting the highest operating frequency. The effect of such stray capacities will be predicted in the following. However, other factors not considered, which cause trouble at high frequencies, such as lead and tube inductances, cathode-to-ground capacitances, etc., will tend to reduce the predicted values. Nevertheless, it is believed that by careful circuit design and choice of suitable tubes, the predicted limits can be closely preserved.

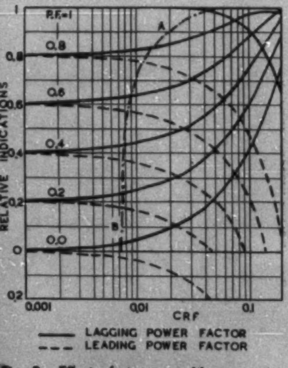


Fig. 8. Effect of stray capacities on meter readings (theoretical)

The presence of a stray capacity C across the output resistance R of the square-wave generator will cause a gradual exponential building up and decay of the voltage, instead of an instantaneous one. With the square-wave voltage applied alone, and with no a-c signal, it can be shown that the average tube currents I_1 and I_2 are given by

$$I_1=I_2=\frac{I_0}{2}\left[1+2CRf\left(\ln\frac{V_c}{E-V_c}-\frac{E}{V_c}\ln\frac{E}{E-V_c}+\frac{E}{V_c}\epsilon^{-\frac{1}{2CRf}}\right)\right]$$

where

f = frequency
V_c = cut-off voltage of the tube
E = amplitude of the square-wave voltage

If, as usual, E is made $2V_c$, then

$$I_1=I_2=\frac{I_0}{2}\left[1+4CRf\left(\epsilon^{-\frac{1}{2CRf}}-0.6931\right)\right]$$

The zero setting is thus not affected by the increase of frequency or stray capacity, but the average current of each tube decreases. It becomes 86% of its value at low frequencies when CRf is increased to 0.05.

When a small a-c signal of the form $F\sin(\omega t-\varphi)$ is applied, the instrument indication can be shown to be

$$I=\frac{S}{\pi}F\left[\cos\left(\varphi-2\pi CRf\ln\frac{E}{E-V_c}\right)+\cos\left(\varphi-2\pi CRf\ln\frac{E}{V_c}\right)\right] \quad (6)$$

and when $E=2V_c$, this becomes

$$I=\frac{2SF}{\pi}\cos(\varphi-4.354CRf) \quad (7)$$

hence

$$\text{relative error}=1-\frac{\cos(\varphi-4.354CRf)}{\cos\varphi}$$

It is seen that the error is independent of the signal amplitude, but depends on the value of the product CRf. Fig. 8 gives the instrument indication at different pf as CRf increases, calculated from equation 7. The curve AB gives the limiting values of CRf if the error is not to exceed 3% of full scale. This gives a value of $CRf=0.0565$ at unity pf and 0.007 at zero pf. With acorn triodes in a cathode-follower circuit, and with proper layout, the value of C can be kept as low as, say, 10 micromicrofarads. Since R is of the order of 1,000 ohms, the instrument is suitable for use at frequencies as high as 5.65 mc at unity pf and 0.7 mc at zero pf. However, equation 6 shows that these values can be increased by increasing E.

Suppressor-Controlled Commutating PF Meter

In the suppressor-controlled commutating pf meter, the square wave is impressed on the suppressor grid of a pentode. When the suppressor grid is driven negative, the anode current falls. The screen-grid current rises, and may attain values in excess of the safe limit. Therefore the developed circuit has in view two purposes: to limit the maximum value of the screen-grid current by lowering the screen potential when the plate current is suppressed, and to make use of the rise of the screen current in increasing the sensitivity. The circuit developed to achieve these two functions is shown in Fig. 9. The screen grid of each tube is cross-connected to the plate of the other tube. Now, when one tube is commutating, the other is suppressed. The plate current of the commutating tube, as well as the relatively high screen current of the suppressed tube, both flowing in the plate-load resistor of the commutating tube, will cause a large voltage drop, so that the voltage at the screen grid of the suppressed tube is largely reduced. The commutating tube, however, has its screen grid connected to the plate of the suppressed tube. Only the relatively small screen current of the commutating tube flows in the plate resistor of the suppressed tube, so that the voltage at the screen grid of the commutating tube is only slightly less than the full d-c supply voltage of the circuit, and thus almost the normal plate current of the commutating tube is obtained. These operations are reversible in each half-cycle of the square wave, so that the suppressed tube will always have a low screen-grid potential and the commutating tube a high screen-grid potential. By making use of the screen-grid currents, an increase of more than 50% in the sensitivity of the instrument has been obtained.

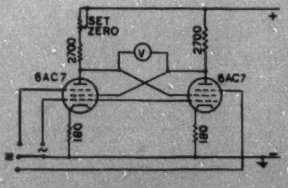


Fig. 9. The suppressor-controlled I cos φ meter

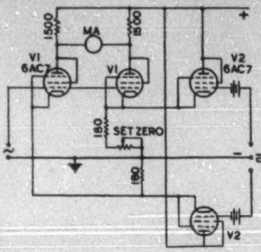


Fig. 10. The cathode-controlled I cos φ meter

The square-wave generator used with this type is essentially the same as that shown in Fig. 5, with minor modifications which mainly imply the addition of a push-pull amplifier stage at the output to cope with the required higher amplitude of the square wave. Mathematical analysis of the circuit is carried out on almost the same lines as for the grid-controlled type, and yields almost similar results.

Cathode-Controlled Commutating PF Meter

The main feature of the cathode-controlled commutating pf meter is that it allows the use of small a-c and square-wave signals, while these are applied to different tube electrodes to allow a suitable grounding point of the system.

Fig. 10 shows the circuit developed for push-pull operation. A common cathode resistor for each pair of tubes *V*1 and *V*2 is used. Tubes *V*2 are biased to cutoff, and a positive square-wave signal is applied to them. In the positive half-cycles of the square wave, the voltage drop across the common cathode resistor should be sufficient to bring *V*1 to cutoff, with the maximum a-c signal applied.

Since the required square wave is of small amplitude, the output stage of the square-wave generator is made a cathode-follower circuit with a small output resistance. This modification makes the square-wave generator suitable for use at much higher frequencies. If it is considered, as mentioned before, that the upper limit of the value of *CRf* is 0.0565 at unity pf and 0.007 at zero pf, and taking *R* as 200 ohms and *C* about 10 micromicrofarads (for a high-frequency pentode in a cathode follower circuit), then the predicted higher limits of the frequency range will be in the neighborhood of 30 mc and 3.5 mc respectively. However, 30 mc seems to be a high value, especially as regards the operation of the square-wave generator (as other factors not considered will cause trouble at such a high frequency), and 10 mc can be considered as the probable higher limit.

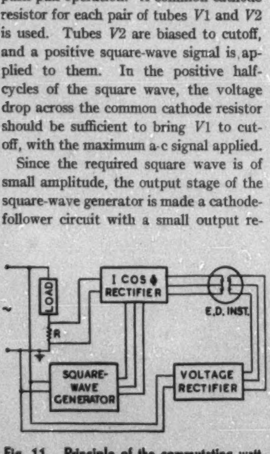


Fig. 11. Principle of the commutating wattmeter

As far as the different sources of error are concerned, this type of pf meter will behave the same as the other two types, and the errors will also be negligible.

Commutating Wattmeter

By means of any of the three aforementioned devices, a direct current proportional to *I* cos φ could be produced. If, now, the voltage across the load is rectified so that a d-c output proportional to the voltage is obtained, the product *VI* cos φ is most easily secured through the use of an ordinary electrodynamic instrument, and the whole arrangement takes the form shown in Fig. 11.

ELECTRODYNAMIC INSTRUMENT

The multiplication of the two d-c signals *I* cos φ and *V* can be effected electronically. However, since the two signals are already obtained in the form of direct currents, it is believed that an electrodynamic instrument would provide a simple and the most straightforward solution. Its design and manufacture would encounter no special difficulties, especially as regards sensitivity and accuracy.

A survey of a number of typical electrodynamic instruments of the laboratory type has shown that in order to obtain full-scale deflection, a relatively large power, of the order of a few hundred milliwatts, is required for the current as well as for the voltage coils.

It should be pointed out that some of the modern types of instruments have better sensitivity, and much more sensitive electrodynamic instruments could be obtained by abandoning the classical technique of using air-core instruments, and using iron-core instruments instead, and consequently a sensitivity as good as that of the moving-coil d-c instrument could be obtained. Iron has many defects, especially when used as a core for a coil carrying direct current to produce a magnetic flux. However, several papers[10-12] have described iron-core instruments which overcome these defects. In the present work, it was not possible to obtain or to construct such a new type of instrument, and therefore stress was given to obtaining as much power from the circuits as possible through the use of normal, powerful receiving-type tubes and comparatively large a-c signals, and sacrificing the use of acorn tubes to extend the working frequency range.

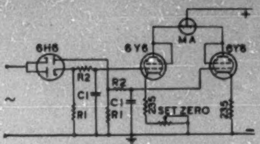


Fig. 12. The voltage rectifier

The usually available electrodynamic instruments have their coils designed for operation in ordinary power supply circuits, and will thus not be suitable for use in conjunction with the circuits under consideration. Therefore the coils of one of these instruments had to be rewound with the proper number of turns and wire diameter and to the proper value of their resistance, but at the same time having almost the same physical shape and size as the original coils. Moreover, the use of push-pull circuits necessitates that each coil should consist of two sections which should be electrically and physically identical, and therefore the bifilar construction was adopted.

The rewound moving coils of the instrument had 1,400 turns and about 1,500 ohms each, while the field coils had 4,000 turns and 900 ohms each. The instrument was provided with a uniform scale.

VOLTAGE RECTIFIER

The voltage rectifier circuit is shown in Fig. 12. A type-*6H6* tube is used for rec-

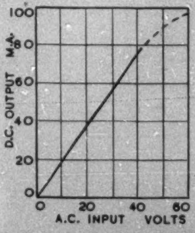


Fig. 13. Characteristic of the voltage rectifier

Vol.
Pt. 1
1955

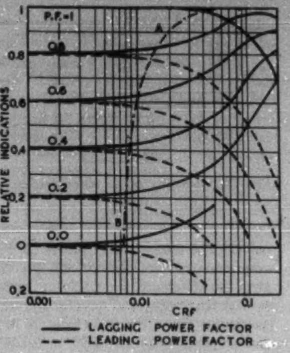
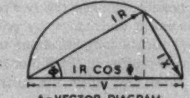


Fig. 14. Effect of stray capacities on meter readings (experimental)

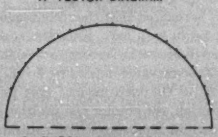
tification. The direct voltage appearing across capacitor *C1* will be equal to the peak value of the applied alternating voltage. If this voltage contains appreciable harmonics, it would be preferable to obtain the average value, which is less sensitive to harmonics, across *C*1. This can be effected through reducing *R*1 and increasing *R*2.

Two type-*6Y6* beam tetrodes, used as cathode followers, are operated in push-pull in order to eliminate errors resulting from nonlinearity that may exist in the working range of their characteristics. These tubes are chosen because of their large current ratings and large mutual conductances, in order to provide the large current necessary for the operation of the electrodynamic indicating instrument. The value of the cathode resistor was chosen with a view to obtaining as large a linear range of characteristic as possible, as maximum an output power as possible, and the linear part of the characteristic extending equally on both sides of the zero input voltage axis so that the tubes are operated without grid bias supply.

The circuit was tested for varying a-c input voltages, the results of which are given in Fig. 13, which shows good linearity from zero up to 45 volts a-c with these tubes. It is obvious that this range can be easily extended as desired by applying only a known fraction of the rectified voltage to the grids of the type-*6Y6* tubes.



A-VECTOR DIAGRAM



B- EXPERIMENTAL RESULTS

Fig. 15. Testing the I cos φ meter, using resistance-capacitance circuit

Experimental Verification

Each type of the pf meters was tested by comparing its readings with those of a dynamometer wattmeter at 50 cycles. Line voltage was applied to the voltage coils of the dynamometer wattmeter as well as to the input of the square-wave generator. This same voltage was made to supply current to a resistive load, and the phase angle of this current was varied by means of a phase shifter introduced between the supply and the load. The same current was passed through the current coils of the electrodynamic wattmeter, while a small fraction of the voltage across the resistive load was applied to the input of the pf meter. Tests were carried for different values of supply voltage, load current, a-c input signal, and for pf from unity to zero. In all cases, results showed close linear relation between the readings of the pf meter and the readings of the electrodynamic wattmeter (divided by the supply voltage). The accuracy as well as the sensitivity were quite satisfactory. Full-scale reading could be obtained by an a-c signal of only a small fraction of a volt, at unity pf, depending on the sensitivity of the indicating instrument. With the grid-controlled meter, an a-c signal of 0.2 volt gave a reading of 24 microamperes (μA) on a bridge-connected indicating instrument of 5,000-ohm resistance, and 43 μA on an instrument of 1,667-ohm resistance. With the suppressor-controlled meter, the same signal gave a reading of 720 μA when using the differential connection, and 300 μA when using a bridge-connected instrument of 1,200-ohm resistance. With the cathode-controlled meter, the same signal gave 120 μA on a bridge-connected 5,000-ohm instrument. The sensitivity can still be increased by using tubes of higher transfer characteristic slopes, or by using more sensitive indicating instruments.

The frequency characteristics of the pf meter, predicted in Fig. 8, were checked simply by increasing the value of *C* instead of increasing *f*, so as to be able to compare the readings to those of the electrodynamic wattmeter. Two similar decade capacitors were connected across the square-wave generator output. The readings obtained by the grid-controlled pf meter are shown in Fig. 14 for different pf. A large input signal (7.5 volts) was used. It is interesting to note that these curves agree quite well in shape and in magnitude with the curves of Fig. 8. The same test was repeated for different input signals, and the error was found to be almost independent of the signal amplitude.

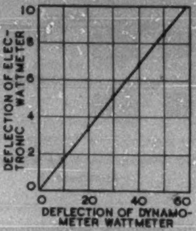


Fig. 16. Results of a test on the commutating wattmeter

Some tests were also carried out on the pf meters at frequencies other than the power-line frequency. In one of these tests an oscillator was connected to a simple series circuit consisting of a noninductive resistance and a calibrated decade capacitor. The frequency and output voltage of the oscillator were kept constant, while the capacitor was varied, and the meter was connected to give readings of *IR* cos φ and *IR*, from which *R*, *X* and cos φ were calculated, and the results plotted in vector form as shown in Fig. 15, which shows also the theoretical semicircle to be expected, for comparison. The figure shows close accuracy of the readings, as the error was always less than 1%. The fact that the experimental points tend to lie outside the circle indicates that this small error is mainly the result of stray capacities. Actually this test is a rather severe one, since a small error in the readings appears exaggerated on the vector diagram.

The tests on the complete wattmeter were carried out on lines similar to those done on the pf meters, i.e., by comparison to the readings of an ordinary electrodynamic wattmeter at power-line frequency. In each test, two of the three variable quantities, viz., alternating voltage to the pf meter, alternating voltage to the rectifier circuit, and pf, were kept constant, and the third varied over its full range. The pf was varied from unity down to zero. The instrument indications are plotted against the electrodynamic wattmeter readings in

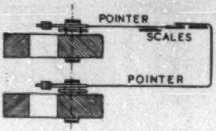


Fig. 17. Arrangement of the twin logarithmic instrument

Fig. 16, using the suppressor-controlled pf meter. The figure shows very good linearity and accuracy. As regards the sensitivity of the wattmeter, it is worth mentioning that, even by using the cumbersome and insensitive electrodynamic multiplying instrument which could be provided, it was possible to obtain full-scale deflection of the instrument by using a current signal of about 1 volt and a voltage signal of about 45 volts.

D-C Division Device

In the preceding sections were described circuits developed by means of which it is possible to measure the inphase component of current, i.e., the product $I \cos \varphi$. The value of the pf, $\cos \varphi$, can be obtained either by measuring the value of I (by the use of the same instrument) and dividing the two readings, or by setting the a-c input signal to the instrument to a predetermined value and calibrating the instrument to give a direct reading of $\cos \varphi$. In both cases, the pf meter can hardly be considered a direct-reading one.

A direct-reading instrument, giving $\cos \varphi$ rather than $I \cos \varphi$, would certainly be much preferable and more useful. Such an instrument would incorporate a device which is capable of dividing two quantities and giving a direct reading of the quotient. Since the quantity $\cos \varphi$ is obtained in the form of a direct current, it is necessary that the other quantity I also be converted to a direct current, and the required dividing device be essentially capable of dealing with d-c signals. A simple d-c division device is described in the following.

The device primarily makes use of the known mathematical relation, that the logarithm of the quotient of two quantities is equal to the difference between the logarithms of the two quantities. As the quantities concerned in this case are direct currents, their logarithmic functions can be simply obtained by the use of log-scale d-c instruments.

Log-scale indicating d-c instruments[13] have actually been designed and manufactured by mass-production technique, with no difficulties from scale distribution. Essentially it is a moving-coil d-c instrument with a magnetic structure of special design. In one case, a full-scale deflection has been obtained with a current of only 150 μA.

The scale of the instrument can be regarded as a part of a slide rule, with the pointer performing the computation. Imagine one scale on which two pointers move, beginning from the same zero point, one pointer deflection being a log function of a certain quantity P, and the other pointer deflection a log function (to the same base of the system of logarithms) of another quantity Q. The angle between the two pointers is then proportional to the logarithm of the quotient P/Q, and the number of scale divisions between the two pointers gives the quotient directly. That the instrument may be direct-reading, means should be available so that this number of scale divisions may be read directly. In other words, the zero reading should now be placed just at the position of one of the pointers, whatever this position may be. This at once suggests that the scale on which the readings are to be taken shall be movable, and fixed to the pointer at its zero-reading point. The other pointer moves on the same scale, and its position will directly give the difference displacement angle, so that its reading will directly give the quotient. Since the movement of the second pointer can be larger or smaller than that of the first, depending on whether the quotient is greater or less than 1, then the movable scale should extend on both sides of its zero-reading point. If the indications are to be read directly as quotients, the indication at this point should read 1. The right-hand side of the scale carries numbers greater than 1, while the left-hand side carries fractions.

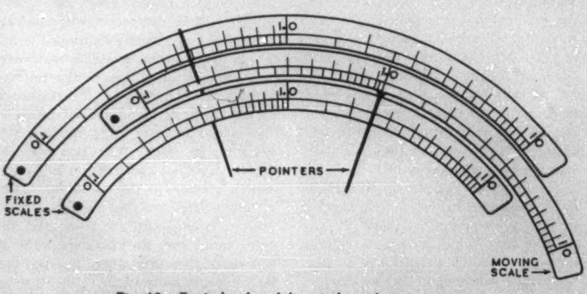


Fig. 18. Typical scales of the twin logarithmic instrument

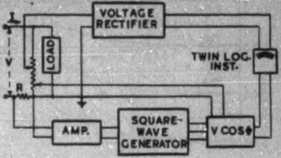


Fig. 19. Principle of the direct-reading pf meter

This principle can be easily effected by means of a twin instrument having two separate and identical log-scale movements. The arrangement of the instrument is shown in Fig. 17. The two movements are placed one above the other and adjusted such that they rotate about the same vertical axis, having the same point of beginning, and are then fixed firmly in their positions. The pointer of the upper movement carries the moving scale, which has to be as light and well balanced as possible, while the pointer of the bottom movement is bent so that it can read on the moving scale.

The twin instrument can be used to give readings of not only their quotient, but of the two quantities which are being divided, by providing each movement with a fixed log-scale. The arrangement of the three scales is shown in Fig. 18.

It is possible to make the two parts of the instrument multirange, by providing each coil with a variable shunt to vary the current passing in the coil in steps of known ratios, by which the instrument readings are to be multiplied.

It is realized that the scale readings do not begin with zero, since the scales are logarithmic. The readings would then begin at some value depending on the sensitivity of the instrument. Not to cause confusion, the pointer rest position (when no current is passing in the instrument) is preferably shifted a little to the left of the start of the scale readings. No readings should be taken when either pointer is in its blank region, in which case the sensitivity of the instrument has to be increased by varying its shunt.

The instrument will not give readings when a negative current passes in it, and therefore it has to be provided with a reversing switch, whose position will indicate whether the reading is to be considered positive or negative.

Finally, it is worth mentioning that the sensitivity of such an instrument would approximate the sensitivity of a conventional linear-scale d-c instrument, provided it would be taken into consideration that the log-scale instrument usually requires more power for equal full-scale rating. The accuracy of the instrument is of the same order as for the linear-scale instrument, or even better accuracy can be obtained in the lower half of the scale.

Direct-Reading PF Meter

The direct-reading pf meter consists of:

1. A commutating circuit to give the inphase component $V \cos \varphi$.
2. A voltage rectifier circuit.
3. A d-c dividing device.

The reason for obtaining $\cos \varphi$ through dividing $V \cos \varphi$ by V (and not dividing $I \cos \varphi$ by I) will be obvious if it is remembered that it is always an essential requirement for the signal obtained from the current element (by the insertion of a resistance in series with the circuit) to be as small as possible. When this small signal is rectified, the direct current obtained might be insufficient for the dividing instrument. Amplification of the signal would therefore be necessary. If the pf meter is to cover a wide frequency range, the amplifier has to be capable of providing a constant gain all over this range, which is not an easy matter, if it could be realized at all.

A direct current which is proportional to the product $V \cos \varphi$ can be obtained in the same manner as already explained for $I \cos \varphi$, i.e., by the use of a square-wave generator operated from the current signal, together with a commutating circuit for which the input is derived from the voltage circuit. An amplifier stage may then be required at the input to the square-wave generator, but this represents no difficulty, since its gain need not be constant, but it should give a negligible phase shift.

A block schematic diagram of the direct-reading pf meter is shown in Fig. 19. The two d-c outputs of the commutating circuit and the voltage rectifier are connected to the twin logarithmic instrument. The instrument then gives direct readings of $V \cos \varphi$, V, and $\cos \varphi$ simultaneously.

Applications of Instruments

It is to be gathered from the foregoing that one instrument can be used for the measurement of voltage, current, pf, or power. This can be readily accomplished through the addition of simple switching arrangements to switch over one or the other of the two a-c signals to the inputs of the square-wave generator, commutating circuit, or voltage rectifier, and to take readings either on a d-c instrument, an electrodynamic indicating instrument, or a twin logarithmic instrument.

The instrument may also be used as a multirange one. The a-c signal derived from the load-current circuit can be changed by making the series resistor vary in known steps. The a-c signal derived from the load-voltage circuit can be changed either by using a voltage divider to take known fractions of this voltage (though this introduces some loss and may cause some difficulties at the higher frequencies), or preferably by using a capacitor voltage divider circuit for the rectified output of the type-*6H6* tube in the voltage rectifying circuit. The instrument can also be used for impedance measurements under actual operating conditions of voltage and frequency. Finally, the instrument can be made as one self-contained a-c operated unit incorporating its power supply circuit.

References

1. An Electron Tube Wattmeter and Voltmeter and a Phase Shifting Bridge, **H. M. Turner, F. T. McNamara.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 18, 1930, pp. 1743–47.

2. Power Measurements with Valves, **H. Lange.** *Archiv für Elektrotechnik*, Berlin, Germany, vol. 26, 1932, pp. 570–79.

3. Valve Wattmeter, **E. Mallett.** *Journal* Institution of Electrical Engineers, London, England, vol. 73, 1933, pp. 295–302.

4. Electronic Wattmeter, **L. R. Malling.** *Electronics*, New York, N. Y., vol. 18, 1945, pp. 133–35.

5. An Improved Electrical Device for Effecting the Multiplication or Division of Independent Quantities, **M. El-Shishini, M. A. El-Said.** British Patent No. *585,936*, Feb. 28, 1947; United States Patent No. *2,486,068*, Oct. 25, 1949.

6. A Thermionic-Tube Measuring Instrument, **Tom B. Wagner.** *AIEE Transactions* (*Electrical Engineering*), vol. 53, Dec. 1934, pp. 1621–23.

7. Proposed Wattmeter Using Multi-Electrode Valves, **J. R. Pierce.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 24, 1936, pp. 577–83.

8. Thermionic Wattmeter, **R. J. Wey.** *Wireless Engineer*, London, England, vol. 14, Sept. 1937, pp. 490–95; Oct., pp. 552–55.

9. Novel Multiplying Circuits with Application to Electronic Wattmeters, **M. A. H. El-Said.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 37, 1949, pp. 1003–15.

10. Tests by the Compensation Method of an A-C Wattmeter Containing a Closed Iron Circuit, **W. Geyger.** *Archiv für Elektrotechnik*, Berlin, Germany, vol. 22, 1929, pp. 119–40.

11. Measurements with the Iron-Cored Dynamometer, **O. Zwierina.** *Elektrotechnik und Maschinenbau*, Vienna, Austria, vol. 53, 1935, pp. 589–97.

12. An Improved Type of D-C Wattmeter of the Shunted Type, **P. MacGahan.** *AIEE Transactions* (*Electrical Engineering*), vol. 60, Jan. 1941, pp. 4–7.

13. Principles of Design of Log-Scale D-C Indicating Instruments, **Allen Stimson, C. F. Taylor.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1011–15.

◆

Discussion

James E. Van Ness (Northwestern University, Evanston, Ill.): This a fine discussion on the use of a commutating circuit for determining components of voltage or current, and on the errors that can arise. I have used this type of circuit in a special-purpose computor with a network analyzer and have found it very satisfactory. The work has been described elsewhere.[1]

In my work, I used a slightly different type of circuit in order to avoid having a steady-state d-c current to balance out. As mentioned in the paper, if the push-pull type of circuit is not used, any variation in the power supply voltage will cause an error in the reading of the indicating meter. Even with the push-pull arrangement, any change in the emissions of the tubes, as from aging or change of heater voltage, will cause an error unless the change is identical in the two tubes. Therefore, difficulties of the type found in the design of d-c amplifiers are likely to occur. I wonder if it would not be well to provide a simple calibrating circuit with this type of instrument for use in zeroing the meter.

For lower frequencies a very simple instrument could be built using a chopper of the type commonly found in servo amplifiers. The chopper could be driven directly by the voltage signal, and the current applied to a d-c meter through the chopper contacts. While the arrangement is not as flexible as one using electronic tubes, it is relatively simple and would avoid any difficulties arising from a steady-state d-c component.

There is one possible source of error in the wattmeter proposed in this paper which I feel should be pointed out. If the rectified voltage signal is not well filtered, the harmonics in its output will react with the harmonics in the output of the commutating circuit and change the reading of the meter. As the harmonic content of the commutated signal can often be much larger than its d-c component, this could be serious. Therefore, the voltage rectifier must include adequate filtering of the signal.

REFERENCE

1. SYNCHRONOUS MACHINE ANALOGUES FOR USE WITH THE NETWORK ANALYZER, James E. Van Ness. *AIEE Transactions*, vol. 73, pt. III, Oct. 1954, pp. 1054–60.

M. Abdel-Halim Ahmed: Mr. Van Ness refers to a simple instrument using a chopper. I have used a similar device for many years in measuring the errors of instrument transformers. The device consisted essentially of a synchronous electromechanical vibrator. It is obvious that such an instrument is suitable only at power frequencies, and has to be retuned whenever the frequency is varied. The troubles which usually arise when vibrating contacts are used are obvious.

To avoid having a steady-state direct current to balance out (although, as will be seen, this does not represent any difficulty) it might be possible to replace the tubes by metal rectifiers, as is usual with modulators used in carrier-current telephone work. This arrangement is still not as flexible and has not as wide a frequency range as one using electronic tubes, but it is definitely much superior to an electromechanical device.

Mr. Van Ness points out that errors may arise from supply voltage variations and changes in the emissions of the tubes when the change is not identical in both tubes, and asks for a calibrating circuit for zeroing the meter. These errors may be either zero-reading errors or errors in the reading of the instrument. With the cathode feedback type of circuit adopted, the equation

S = slope of the transfer characteristic of either tube

$$=\frac{\text{cathode current}}{\text{grid voltage}}=\frac{\mu}{\mu+1}\,\frac{1}{R_k+\dfrac{\rho+R_a}{\mu+1}}$$

shows that the tube current is almost independent of the tube parameters μ and ρ, since the fraction $\mu/(\mu+1)$ is almost equal to unity and $(\rho+Ra)/(\mu+1)$ is made much smaller than R_k. The tube current, as well as the value of S are therefore determined by the value of the cathode resistor R_k. This value is adjusted to be equal for both tubes by the small variable resistance added in the cathode circuit of one of the tubes, as shown in Figs. 3, 9, and 10. The adjustment of this resistance for zero no-signal current in the indicating instrument will also make the transfer characteristics of both tubes identical. In this case, trouble with the zero reading is entirely eliminated, and the error in the instrument reading due to supply voltage variations, aging or replacement of the tubes, etc., is reduced to a negligible amount.

However, the instrument can always be readily calibrated by switching it to read either current or voltage, and comparing its reading with that of an ordinary ammeter or voltmeter, and there will be no need to add a special calibrating circuit. If any error is noted, this will only involve multiplying the instrument reading by a constant factor.

I agree that for using the instrument as a wattmeter, the rectified voltage signal has to be well filtered. The voltage rectifier circuit takes care of this. However, as mentioned in the paper although not shown in the figures, the indicating instrument in the commutating circuits may be conveniently shunted with a capacitor to by-pass the alternating components in the output of the commutating circuit; and the same procedure can be followed with the voltage rectifier circuit. No error in the wattmeter reading will result in this case.

Theory of Magnetic Cross Valves

W. H. HIGA
NONMEMBER AIEE

A PHYSICAL and mathematical theory of the interaction of two mutually orthogonal coils in a ferromagnetic medium is developed. For purposes of analysis, an idealized configuration with cross-sectional view, as shown in Fig. 1, is studied. Two coils, x and y, are wound at right angles to each other about a cylinder of ferromagnetic material of permeability μ. The structure is then immersed in a medium of infinite permeability. This last requirement is not stringent, it being required only that the saturation effects be confined to the central core for purposes of analysis. It is assumed further that the core material has no hysteresis or eddy-current losses while its typically nonlinear magnetization curve is as shown in Fig. 2. It is shown how the nonlinearity in the magnetization curve results in the interaction of the orthogonal circuits.

A practical version of the configuration under analysis was first described by H. J. McCreary who has made many interesting applications of the device.[1] The primary importance of the magnetic cross valve is in frequency-changing applications. Large amounts of power can be supplied at one-half or one-third line frequency.

The nature of nonlinear coupling is such that there must exist, initially, a field component orthogonal to the applied modulating field. It is shown that, in the absence of the orthogonal field, there can be no interaction. Although, in the first example, a constant orthogonal field is used, it is also possible to have a periodically varying field modulated by the applied one. This type of operation requires a capacitor in the secondary circuit. An oscillatory state may then be established and maintained in the secondary circuit. This phenomenon is closely related to ferroresonance in other types of circuits containing iron-cored inductors and capacitors.[2] The importance of ferroresonance lies in the fact that the sustained oscillations may be at a multiple or submultiple of the driving generator frequency.

Physical Analysis

To explain the physical mechanism through which interaction can occur in orthogonal circuits, it will be convenient to assume that the magnetic medium is homogeneous and isotropic. Quantities are defined as follows:

H_x = magnetic field caused by coil x
H_y = magnetic field caused by coil y
φ_x = flux-threading coil x
φ_y = flux-threading coil y
$\mathbf{H}$ = field vector with components H_x and H_y; $H=|\mathbf{H}|$
$\boldsymbol{\varphi}$ = flux vector with components φ_x and φ_y; $\varphi=|\boldsymbol{\varphi}|$
$\dot{\varphi}=d\varphi/dt$
$\ddot{\varphi}=d^2\varphi/dt^2$

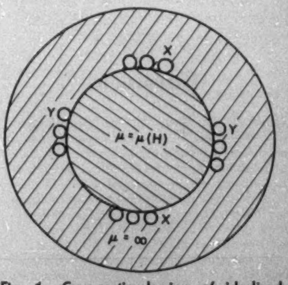


Fig. 1. Cross-sectional view of idealized magnetic cross valve

---

Paper 55-8, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted March 18, 1954; made available for printing October 27, 1954.

W. H. HIGA is with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif., and was with North American Aviation, Inc., Downey, Calif., when the paper was prepared.

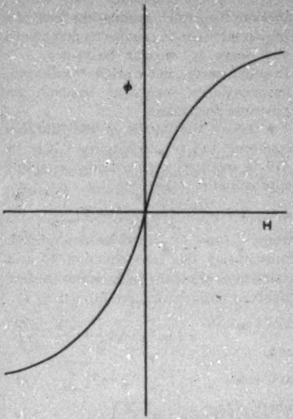


Fig. 2. Magnetization curve for core of Fig. 1

Fig. 3 shows a typical plot of the loci for the **H** and φ. Here it is assumed that H_x is held constant and H_y is varied; furthermore, the permeability is assumed to be constant. It can be seen readily that in the case of constant permeability, the loci for **H** and φ can be brought into coincidence by a proper choice of scale factor, There if, of course, no interaction in this case.

Fig. 4 shows the loci for **H** and φ when the nonlinearity of the core material is taken into account. In particular, a constant field component H_x is assumed and H_y is varied. If the component H_x is sufficiently large, there will be little change in the length of φ with changes in H_y; however, φ will always be parallel to **H**. Thus, it can be seen that the loci for **H** and φ cannot be brought into coincidence here. A study of Fig. 4 shows that a variation in H_y causes a variation in the component φ_x. Thus, if H_y varies periodically with time, there results a periodic variation φ_x in which induces an electromotive force in coil x. Fig. 5 shows a plot of H_y and φ_x as functions of time. It is apparent that the variations in φ_x occur at twice the frequency of the variations of H_y. Thus, a frequency doubler is easily obtained through the connections shown in Fig. 6.

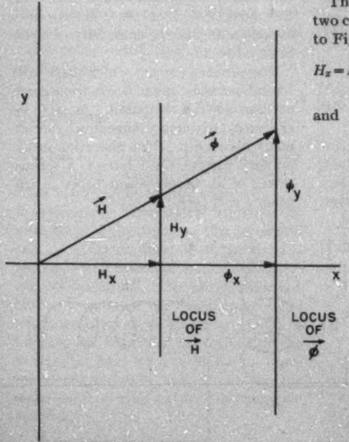


Fig. 3 (left). Loci for H and φ when μ=constant

Mathematical Analyses

The basic assumption in this study is that the phenomena associated with the magnetic cross valve are phenomena on the macroscopic scale. That it to say, it is assumed that the total flux vector φ and an average field vector **H** are adequate in deducing a theory. The actual distribution of the field in the core need not be discussed. This assumption was, of course, made in the foregoing.

The nonlinear magnetization curve of Fig. 2 may be expressed analytically as

$$H=a\varphi+b\varphi^3+c\varphi^5+\ldots \tag{1}$$

where the coefficients a, b, c, etc., are obtained from the curve. It will be adequate here to drop all but the first two terms. Hence, the basic equation is

$$H=(a+b\varphi^2)\varphi \tag{2}$$

The components associated with the two coils are readily obtained by referring to Fig. 4 and are

$$H_x=H\cos\theta=H(\varphi_x/\varphi)=a\varphi_x+b(\varphi_x^2+\varphi_y^2)\varphi_x \tag{3A}$$

and

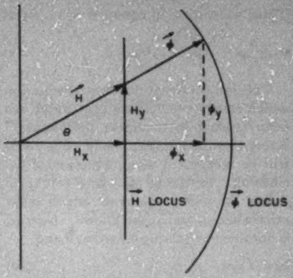


Fig. 4. Loci for H and φ when $\mu=\mu(H)$

$$H_y=H\sin\theta=H(\varphi_y/\varphi)=a\varphi_y+b(\varphi_x^2+\varphi_y^2)\varphi_y \tag{3B}$$

Before deducing from equations 3, the basic equations of ferroresonance, it will be convenient to normalize the quantities H_x, H_y, φ_x, φ_y as follows

$$H_x=ki_x=i_x \tag{4A}$$

$$H_y=ki_y=i_y \tag{4B}$$

and

$$\varphi_x=Ke_x=e_x \tag{5A}$$

$$\varphi_y=Ke_y=e_y \tag{5B}$$

The same number of turns in the primary coil y and the secondary coil x are assumed, and constants of proportionality k and K are assumed equal to unity. Thus, equations 3 can be written as

$$i_x=a\varphi_x+b(\varphi_x^2+\varphi_y^2)\varphi_x \tag{6A}$$

$$i_y=a\varphi_y+b(\varphi_x^2+\varphi_y^2)\varphi_y \tag{6B}$$

Consider now the circuit shown in Fig. 7; the secondary current is found from

$$\varphi_x+\frac{\int i_x dt}{C}=0 \tag{7}$$

or

$$i_x=-C\dot{\varphi}_x \tag{8}$$

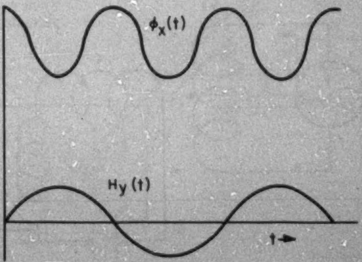


Fig. 5 (right). Variations of $H_y(t)$ and $\varphi_x(t)$ as functions of time

l, 74
1
55

Thus, substution in equation 6(A) yields

$$\ddot{\varphi}_x+\frac{a}{C}\varphi_x+\frac{b}{C}(\varphi_x^2+\varphi_y^2)\varphi_x=0 \tag{9}$$

This is a nonlinear differential equation which closely resembles a type which has been studied in nonlinear mechanics.[3,4] Except for the nonlinear term, equation 9 will be recognized as an equation of oscillation. Equation 6(B) now serves to define the primary current which must exist when φ_x and φ_y are known. Now φ_x is obtained by solving equation 9 and

$$\varphi_y=\int e_y dt \tag{10}$$

To find particular solutions to equation 9 is the next step. It should be pointed out that most nonlinear differential equations do not have general solutions as is the case with linear ones. The equation must be solved for each new set of conditions. This will be illustrated by the examples that follow.

CASE I: THE PHASE-SHIFTING TRANSFORMER

Let $e_y=E_0 \sin \omega t$ be the voltage applied to the primary winding and let $\varphi^2=\varphi_x^2+\varphi_y^2=\Phi_0^2$ (a constant).

Here

$$\varphi_y=-\frac{E_0}{\omega}\cos \omega t=-\Phi_0 \cos \omega t$$

and

$$\varphi_x=\frac{E_0}{\omega}\sin \omega t=\Phi_0 \sin \omega t \tag{11}$$

to satisfy the second condition. Substituting in equation 9

$$-E_0\omega+\frac{a}{C}\frac{E_0}{\omega}+\frac{b}{C}\left(\frac{E_0}{\omega}\right)^3=0 \tag{12}$$

or

$$C=\frac{1}{\omega^2}\left(a+b\frac{E_0^2}{\omega}\right) \tag{13}$$

This results states that if a voltage $E_0 \sin \omega t$ is applied to the primary, a voltage

$$e_x=\dot{\varphi}_x=E_0 \cos \omega t \tag{14}$$

will appear across the secondary, provided that the relation of equation 13 is satisfied and that an initiating transient is introduced in the secondary. The last condition is necessary since interaction cannot occur in the absence of an orthogonal field.

In practice, the condition specified by equation 13 is not as critical as it may first appear. The reason for this is that the constant a and b actually depend on the operating range on the magnetization curve. Hence, when at least close to the proper range, it is possible for the circuit to adjust itself to satisfy equation 13.

It is, in fact, this flexibility in the circuit which permits at least a little power from the secondary circuit to be drawn. When the resistive load becomes too large, the circuit will abruptly stop oscillating. These matters are concerned with stability criteria in ferroresonance and are not discussed here.

CASE II: GENERATION OF SUBHARMONIC OF ORDER ONE-HALF

Let

$$\varphi_y=\Phi_0+\frac{E_0}{\omega}\cos \omega t$$

with

$$\Phi_0>>\varphi_x \text{ or } \varphi_y$$

then equation 9 becomes

$$\ddot{\varphi}_x+\frac{a}{C}\varphi_x+\frac{b}{C}\left[\Phi_0^2+2\Phi_0\frac{E_0}{\omega}\cos \omega t\right]\varphi_x=0 \tag{15}$$

This equation may be written as

$$\ddot{\varphi}_x+(\alpha+2\beta \cos \omega t)\varphi_x=0 \tag{16}$$

where

$$\alpha=\frac{a}{C}+\frac{b}{C}\Phi_0^2 \tag{17}$$

and

$$\beta=\frac{bE_0\Phi_0}{C\omega} \tag{18}$$

Equation 16 is the Mathieu equation which has the interesting feature that, when α and β satisfy certain conditions, periodic solutions obtain. Moreover, this solution which assumes the form of a trigonometric series has for its first term a sine wave of angular frequency $\omega/2$. In other words, under proper conditions, ferroresonance occurs at one-half the generator frequency.

Although the details of the Mathieu equation[5] will not be discussed here, it will be instructive to consider a simple approximation. Let

$$\varphi_x=\lambda_1 \cos(\omega t/2)+\lambda_2 \cos(3\omega t/2) \tag{19}$$

where λ_1 and λ_2 are to be determined. Substituting this into equation 16 and equating coefficients of like terms in time yields

$$a_{11}\lambda_1+a_{12}\lambda_2=0 \tag{20}$$

and

$$a_{21}\lambda_1+a_{22}\lambda_2=0 \tag{21}$$

where

$$a_{11}=\frac{1}{C}\left(a+b\Phi_0^2+\frac{bE_0\Phi_0}{\omega}\right)-\frac{\omega^2}{4}$$

$$a_{12}=a_{21}=\frac{bE_0\Phi_0}{C\omega}$$

and

$$a_{22}=\frac{1}{C}(a+b\Phi_0^2)-\frac{9\omega^2}{4}$$

Equations 20 and 21 can yield unambiguous solutions for λ_1 and λ_2 only if

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=0 \tag{22}$$

Thus, there is again a condition on the circuit parameters which must be satisfied for subharmonic resonance. In obtaining equations 20 and 21, a term ($\cos 5\omega t/2$) has actually been discarded; this must be done unless it is desired to assume more exactly an infinity of terms in equation 19.

The primary current which must exist for subharmonic resonance in the secondary can now be computed. As is to be expected, the primary current will contain many harmonics. This distortion in the primary current wave form is a serious defect of all ferroresonant types of frequency changers.

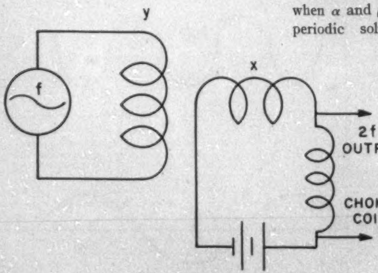
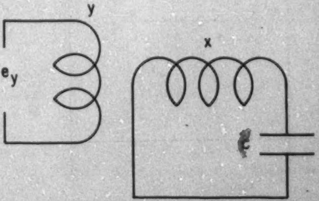


Fig. 6 (left). Simple frequency doubler

Fig. 7 (right). Ferroresonant transformer



Vol. 7
Pt. 1
1955

CASE III: GENERATION OF SUBHARMONICS OF ORDER ONE-THIRD

Let

$$\varphi_y = \Phi_y \sin \omega t$$

and

$$\varphi^3 \neq \text{constant}$$

Here, a generalized problem of which case I was a specialization is being considered. Accordingly

$$\varphi_x = \sum_{n=1}^{\infty} \Phi_{n/k} \sin \frac{n\omega t}{k} \qquad (23)$$

where k is an integer. Substitution of φ_x in to the differential equation leads to

$$\left(\frac{a}{C} - \frac{n^2\omega^2}{k^2} + \frac{b}{C}\Phi_y^2 \sin^2 \omega t\right) \times \sum \Phi_{n/k} \sin \frac{n\omega t}{k} + \left(\sum \Phi_{n/k} \sin \frac{n\omega t}{k}\right)^3 = 0 \qquad (24)$$

The coefficients Φ will have nontrivial solutions only if there are common time factors in the first and second terms of equation 24. Hence, in view of the identity

$$\sin 3\theta = 3 \sin \theta - 4 \sin^3 \theta$$

and the occurrence of the cubed term in equation 24, the following approximate solution can be justified

$$\varphi_x = \Phi_{1/3} \sin \frac{\omega t}{3} + \Phi_1 \sin \omega t \qquad (25)$$

When this solution is substituted into the differential equation, two equations are obtained for the solution of the Φ_s'. As in the previous cases, the circuit parameters must satisfy certain relations in order that ferroresonance of order one-third occurs. The details, unfortunately, are cumbersome with algebra and will be omitted here.

Conclusions

It has been shown that the mechanism of frequency changing in the magnetic cross valve can be readily analyzed. Although many simplifying assumptions were made in the course of this investigation, it appears that all of the experimental observations of H. J. McCreary[1] can be explained. The use of only the first two terms in the expansion for the magnetization curve has the effect of limiting the harmonic distortion for the modes that have been considered. More terms are required to account for the many more complex modes of operation possible.

The differential equation 9 which governs the ferroresonance phenomena is related to the equation of Duffing which has been studied extensively in nonlinear mechanics.[3,4] The inclusion of dissipative currents in the system adds a first time derivative term $\dot{\varphi}$ to equation 9. The principal effect of such a term would be to introduce phase shift into the voltages and currents.

References

1. THE MAGNETIC CROSS VALVE, H. J. McCreary. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1868–75.

2. SUBHARMONICS IN CIRCUITS CONTAINING IRON-CORED REACTORS, I. Travis, C. N. Weygand. *AIEE Transactions*, vol. 57, Aug. 1938, p. 423.

3. NON-LINEAR VIBRATIONS (book), J. J. Stoker. Interscience Publishers, New York, N. Y., 1950.

4. NON-LINEAR MECHANICS (book), N. Minorsky. J. W. Edwards, Ann Arbor, Mich., 1947.

5. THEORY AND APPLICATIONS OF MATHIEU FUNCTIONS (book), N. W. McLachlan. Oxford University Press, London, England, 1947.

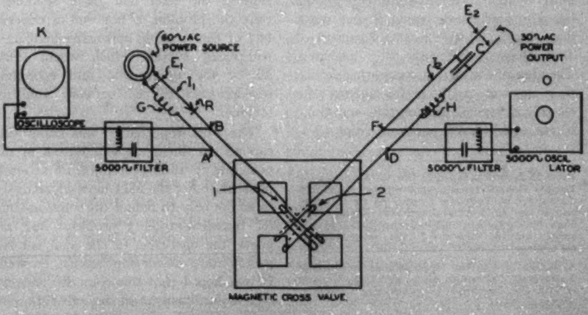
◆

Discussion

Harold J. McCreary (Automatic Electric Company, Chicago, Ill.): This is a very commendable contribution to the literature concerning the magnetic cross valve. While, as the author states, many simplifying assumptions were made, nevertheless he arrives at the Mathieu equation as a solution representing what takes place in the magnetic cross valve. This also appears to be Edson's conclusion in his book *Vacuum Tube Oscillators*.[1]

Since the apparent object of the paper is to clarify by means of mathematics, what takes place in the magnetic cross valve, and why, perhaps I can appropriately add some experimental evidence on the reasons for what takes place. (The process itself was covered fairly well by oscillograms in my article, "The Magnetic Cross Valve," see reference 1 of the paper).

Fig. 8. Auxiliary test equipment



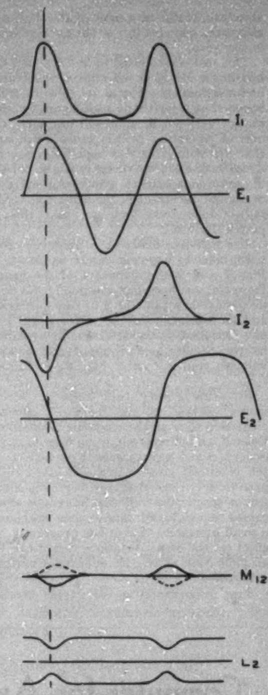


Fig. 9. Oscillograms of input and output voltage and currents

MUTUAL INDUCTANCE

The results can obviously be attributed to the peculiar relationship between the input coil and output coil. This relationship is usually called "mutual inductance" and is assumed to be a constant. In the magnetic cross valve, because of its unusual structure, this is zero most of the time, and alternates periodically; so the average value is also zero.

This test was devised to indicate how the instantaneous value of mutual inductance and self-inductance varies. This test assumes that a relatively high-frequency current of very small value will not affect the normal operation of the cross valve and that the voltage of this high frequency induced in one coil is a direct indication of the instantaneous mutual inductance and rate of change of current in the other coil. The independent tests indicate that this is practically true.

The test was made on a magnetic cross valve used as a power ringer as shown in Figs. 7 and 9 of reference 1 of the paper. However, auxiliary test equipment was connected as shown in Fig. 8 of this discussion and the operation of the power ringer was not disturbed by the attachments as shown by the oscillograms of input and output voltages and currents, Fig. 9, I_1 E_1 and I_2E_2.

By using a very high frequency alternating current, it can be independently channeled through the magnetic cross valve, isolated, and filtered out, using filtering components which have practically no effect on the magnetic cross valve and its normal operation. Also, with such high frequency the magnetization of the magnetic core caused by this high frequency can be of such a small increment that it has practically no effect on the core saturation; yet it can be an indication of the instantaneous mutual inductance.

An oscillator which delivers 5,000 cycles is connected through a 5,000-cycle-per-second filter to the magnetic cross valve coil designated 2, imposing a high frequency on this coil.

Across the coil designated 1 of the magnetic cross valve, oscilloscope K is connected through a filter which passes only the 5,000 cycles. The oscillograph envelope M_{12} is what the oscilloscope K indicated.

The mutual inductance as indicated by the oscillogram envelope is zero most of the time, but shows up by allowing spurts of the 5,000 cycle to pass from coil 2 to coil 1 for very short intervals.

The mutual inductance occurs when the current in both coil 1 and coil 2 are a maximum. Successive couplings by mutual inductance are reversed relative to each other.

Self Inductance

Inductance is always measured by its effect on changing current. In fact, since R is constant for all practical purposes we may use the reactive voltage to indicate the value of X and of L, since X is proportional to L.

I feel therefore that this test indicates how the self-inductance of the magnetic cross valve is changing without seriously modifying the magnetic cross valve action.

An oscillator which delivers 5,000 cycles is connected through a filter to the magnetic cross valve coil designated 2, at F D, imposing a high frequency on this coil. To these same points F and D an oscillograph like K is connected through a filter which passes the 5,000 cycles only. The oscillograph envelope L_2 is what the oscilloscope indicated.

The impedance of the secondary coil as indicated by the envelope oscillogram L_2 is practically a straight line except at the time that mutual inductance comes into effect, throwing the primary and secondary circuits in parallel, and so decreasing the impedance. It can therefore be assumed that the self-inductance is a constant except for the time period when mutual inductance becomes effective.

I conclude that the self-inductance of coil 2 is a constant as shown by the oscillograph envelope L_2 and that the total reactive component of coil 2 is modified principally by the mutual inductance becoming effective and, in effect, connecting the primary circuit in shunt with the secondary circuit.

These tests substantiate our previous theories of how the magnetic cross valve works. In particular, as described in Fig. 2 of reference 1 of the paper.

Reference

1. Vacuum Tube Oscillators (book), W. A. Edson. John Wiley & Sons, Inc., New York, N. Y., 1953, p. 357.

W. H. Higa: The author thanks Mr. McCreary for his interesting discussion. The use of a high frequency carrier to monitor a saturable reactor device is certainly unique and should find many applications.

Computer for Automatizing Network Analyzer Operation

A. KUSKO
MEMBER AIEE

P. N. HELLER
ASSOCIATE MEMBER AIEE

THE recent activity and interest in the development of computer methods for speeding up the solution time for power-system transient-stability problems has prompted the reporting of some work done on the Massachusetts Institute of Technology (MIT) network analyzer. The idea of coupling a computing device to a network analyzer for solving the electromechanical differential equations of the synchronous machine rotor positions following a fault or other large disturbances to the system is certainly not a new one. Many people have considered the idea and have concluded that the problem is not one of concept but of constructing and instrumenting a computing system which would fulfill all of the operational and economic requirements of the network analyzer service.

The work reported in this paper was carried out during an exploration of the possibility of adding computer units to each phase shifter of the MIT network analyzer so as not only to reduce the solution time for transient-stability problems but improve the operation of the analyzer for steady-state power-flow studies as well. It was hoped that this computer system would facilitate work on other power-system problems which heretofore were not undertaken because of computational difficulties.

The Power-System Problem

The solution of a power-system transient-stability problem requires the determination of the rotor angle of each machine or group of machines in the system as a function of time following the initiating disturbance. The rotor angle of a given machine, e.g., the kth machine in an n-machine system, is generally obtained from the following approximate equation which has been derived in the Appendix

$$\delta_k=\delta_k(0)+\omega_k(0)t+\frac{1}{M_k}\int_0^t\int_0^t(P_{ik}-P_{ok})dtdt \tag{1}$$

where

δ_k = the phase angle of the voltage behind the transient reactance of the kth machine

$\delta_k(0)$ and $\omega_k(0)$ = the initial value of δ_k and its initial rate of change respectively

M_k = the inertia constant of the kth machine

P_{ik} = the mechanical power supplied to the kth machine by its prime mover

---

Paper 55-5, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31-February 4, 1955. Manuscript submitted September 7, 1954; made available for printing October 27, 1954.

A. Kusko is with the Massachusetts Institute of Technology, Cambridge, Mass., and P. N. Heller is with the Raytheon Manufacturing Company, Waltham, Mass.

The electric power output P_{ok} of the kth machine is given by

$$P_{ok}=\frac{E_k^2}{Z_{kk}}\cos\theta_{kk}-\sum_{j=1}^{n}\frac{E_kE_j}{Z_{kj}}\cos(\delta_k-\delta_j+\theta_{kj}) \qquad (2)$$

where

$j=1, 2, 3 \ldots k-1, k+1 \ldots n-1, n;\ j\neq k$
$E_k\underline{|\delta_k}$ = the voltage behind the transient reactance of the kth machine
$E_j\underline{|\delta_j}$ = the voltage behind the transient reactance of the jth machine
$Z_{kk}\underline{|\theta_{kk}}$ = the driving-point impedance from the kth machine to the system including the transient reactance of the machine
$Z_{kj}\underline{|\theta_{kj}}$ = the transfer impedance from the kth machine to the jth machine in the system including the transient reactances of the machines

Substitution of equation 2 into equation 1 yields

$$\delta_k=\delta_k(0)+\omega_k(0)t+\frac{1}{M_k}\int_0^t\int_0^t\left(P_{ik}-\frac{E_k^2}{Z_{kk}}\cos\theta_{kk}+\sum_{j=1}^{n}\frac{E_kE_j}{Z_{kj}}\cos(\delta_k-\delta_j+\theta_{kj})\right)dtdt \qquad (3)$$

Similar equations must be written for the remaining $(n-1)$ synchronous machines in the system and the resulting set of n simultaneous differential equations must then be solved. However, since a mathematical solution in closed form exists only in exceptional cases, methods other than those of formal mathematical analysis must be used.

Previous Work

Nearly all multimachine transient-stability problems presently are solved on network analyzers using the step-by-step method. The tediousness and inaccuracy of this method has led to a search for suitable machine methods for solving problems of this type. There have been a number of published accounts of successful designs of computers for this purpose, but no data on the installation and continued use of such computers in daily service have been provided. These computers fall into several categories which will be reviewed here.

In the first of these categories are the large-scale general-purpose analogue or digital computers which are theoretically capable of solving representative multimachine transient-stability problems. However, the transfer impedances and prefault conditions must be obtained first from a network analyzer or by manual calculation and then set on the computer. This requires sufficient time so that, despite the speed and ease with which the actual computation can be carried out, not much is gained by the use of a separate computer for routine multimachine transient-stability studies. A differential analyzer has been used for special work by Concordia and others.[1-3] A synchronous machine was represented in detail to study the effect of the excitation system of this machine and its prime-mover governor on the transient stability of a 2-machine system. An electric analogue computer operating on the same principle has been developed by Boast and Rector.[4] The mechanical parameters of the synchronous machines are represented in this computer by resistors and inductors; transfer admittances of the power system are represented by d-c amplifiers and photoformers. A conventional electronic analogue computer can be used in about the same manner. To solve an n-machine problem by either of these methods, $n(n-1)$ amplifiers and at least 1/2 $n(n-1)$ function generators are required to represent the transfer admittances.[4] Thus, 30 amplifiers and at least 15 function generators are needed for the solution of the electric part of a 6-machine problem.

Another class of computer[5] makes use of small synchronous machines and a 3-phase model representation of the power system under consideration. The d-c motor-driven synchronous machine sets are equipped with voltage and speed regulators, interchangeable rotors, and provisions for changing the field time constant of each machine by means of a feedback network so that both the static and the dynamic characteristics of the machine-prime mover sets in the actual system can be duplicated. Faults are applied cyclically and the resultant disturbance is recorded by means of a cathode-ray oscilloscope. The use of interchangeable rotors in this computer permits the study of the effects of rotor saliency and amortisseur windings, while the 3-phase representation of the power system makes it possible to study the effect of unsymmetrical faults and single-phase switching without resorting to sequence networks. In use, the model system will behave like a real power system and the usual problems of frequency control, load division, and voltage regulation will make its adjustment somewhat difficult.

The third class of computer stems from the original suggestion by Kimbark[6] in 1938 for coupling an analogue computer to a network analyzer and has the sound basis that the network analyzer actually solves the electric part of the electromechanical differential equations while the computer solves the mechanical part. Moreover, the power system representation on the network analyzer for transient-stability studies is practically the same as the representation used for the study of steady-state power flow; the transition from steady-state to transient analysis and back can therefore be made with little disruption. The all-electronic computers, described by Kaneff[7] and Shen and Lissner[8] for use with 10-kc network analyzers, are of this type; each computer unit and phase shifter comprises a single assembly. Van Ness[9] described a similar more elaborate system but gives limited construction details. For network analyzers operating at conventional frequencies Bekey and Schott[10] proposed a computer system in which the components of a mechanical differential analyzer are used to control the rotor position of each existing phase shifter. The use of human operators to read each wattmeter and feed data to the differential analyzer solves the power-sensing problem, although automatic methods are mentioned.

Experimental Computer Unit

A computer system similar in concept to the third class described in the foregoing was designed to work with the MIT network analyzer and an experimental computer unit for such a system was built and tested. This unit positions one phase shifter of the network analyzer in accordance with the solution of the power-system electromechanical differential equation. For a complete installation one such computer unit is required for each phase shifter used in the power-system representation.

The block diagram of the experimental unit is shown in Fig. 1. Watt-hour meter no. 1 is a 2-element watt-hour meter; one element measures the output energy of the phase shifter and the other element measures energy proportional to the mechanical energy supplied to the machine represented by the phase shifter. The number of revolutions made by the watt-hour meter disk, therefore, is proportional to the time integral of $(P_{ik}-P_{ok})$. A photoelectric pickup servomechanism positions a variac so that the variac output voltage V is proportional to the number of revolutions of the watt-hour meter disk. Thus

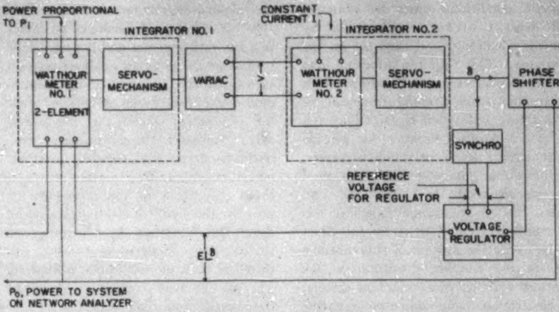


Fig. 1. Block diagram of the experimental computer unit

$$V = V(0) + K_1 \int_0^t (P_{ik} - P_{0k})dt \tag{4}$$

where

V = the output voltage of the variac
$V(0)$ = the initial value of that voltage
K_1 = a constant

The second integration is performed by the single-element watt-hour meter no. 2. One coil of this meter is connected to the output terminals of the variac and the other coil is supplied with a constant current I. Since the phase angle between this constant current I and the variac voltage V is fixed, the disk speed of the watt-hour meter no. 2 is proportional to the product IV, and the change in disk position is proportional to the time integral of IV. The rotor of the phase shifter is positioned by a photoelectric pickup servomechanism so that the angle through which the rotor turns is proportional to the number of revolutions made by the watt-hour meter disk. It follows that

$$\delta = \delta(0) + K_2 I \int_0^t V dt \tag{5}$$

where

δ = the mechanical angle of the phase-shifter rotor
$\delta(0)$ = the initial value of δ
K_2 = a constant
I = the constant current

Combining equations 4 and 5 yields

$$\delta = \delta(0) + K_2 I V(0)t + K_1 K_2 I \int_0^t \int_0^t (P_{ik} - P_{0k}) dt dt \tag{6}$$

By assigning proper values to the constants K_1, K_2, and I in equation 6 for the computer unit, this equation can be made analogous to equation 1 for the system. The solution of equation 2, the power-flow equation, is obtained from the network analyzer and the combination of analyzer and computer unit will, therefore, solve the complete electromechanical differential equation for one machine.

The voltage regulator, shown in Fig. 1, continuously corrects for the internal-impedance drop of the phase shifter in such a manner that the output voltage E of the combination of phase shifter and voltage regulator is constant in magnitude and that the phase angle of this voltage is equal to the computed value δ. The reference voltage for the voltage regulator is obtained from the rotor of a synchro which is coupled through an antibacklash gear train to the output shaft of integrator no. 2. The stator of this synchro is supplied from a source of constant 3-phase voltage. Since the output voltage of the phase shifter itself is nearly equal to the desired value, the voltage regulator need only furnish the small voltage necessary to correct for the internal-impedance drops. An electronic amplifier is used in the regulator of the experimental computer unit but other schemes may be used in permanent installations.

Fig. 2 shows the watt-hour meter of integrator no. 2 with the follow-up elements of the servomechanism. The meter itself is a standard 15-ampere 115-volt 60-cycle-per-second watt-hour meter; the current coil is supplied by a 6.3-volt 50-ampere transformer. Part of the phase shifter which is positioned by the computer unit is shown on the left of the watt-hour meter. This phase shifter, one of the regular units of the MIT network analyzer, is used with only very minor mechanical modifications.

Solution of Sample Problems

Two sample problems were run to test the experimental computer unit. The first test problem was the prediction of the power-angle excursion of a synchronous machine suddenly connected out of phase to an infinite bus. For certain initial conditions, there is a mathematical solution in the form of tabulated eliptic integrals. This solution was used as a convenient means of checking the computer accuracy. After 1 hour of computer operation corresponding to 6 seconds of system time, the amplitude of machine oscillation as given by the computer differed by less than 12 per cent from that computed from the mathematical equation. Some of the sources of this error are known and can probably be eliminated.

The second sample problem was the prediction of the power-angle transient for a single machine following a 3-phase fault and circuit-breaker operation in the transmission network connecting the machine to an infinite bus. The system, shown in Fig 3, was set up on the network analyzer and the power-angle excursions as a function of time were obtained from the computer unit for 3-cycle, 6-cycle, and 9-cycle clearing and variations thereof. Results are shown in Fig. 4. A time scale of 10 minutes of computer operation per second of system time made it possible to switch the fault and line connections on the analyzer manually without stopping the computation.

Use of the Computer Unit as a Power Regulator

Steady-state power-flow studies constitute the major portion of the work done on network analyzers. If the computer units of a permanent installation can be used to control the power output and out-

Fig. 2. Watt-hour meter of integrator no. 2

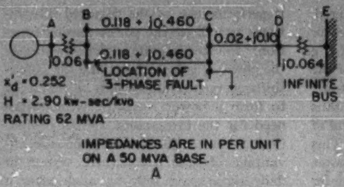


Fig. 3(A) (left). Impedance diagram of the 1-machine power system used for the computer test

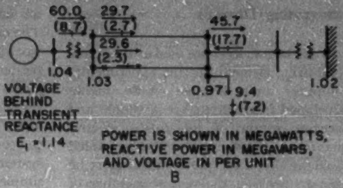

60.0
(8.7)
29.7
(2.7)
29.6
(2.3)
45.7
(17.7)
1.04
1.03
0.97
9.4
(7.2)
1.02
VOLTAGE BEHIND TRANSIENT REACTANCE
E1 = 1.14
POWER IS SHOWN IN MEGAWATTS, REACTIVE POWER IN MEGAVARS, AND VOLTAGE IN PER UNIT
B

Fig. 3(B) (right). Power-flow diagram of the system for the prefault condition

put voltage of the phase shifters in the analyzer, a saving in the time required to adjust the phase shifters can be realized. This can be done by increasing the computing speed of the computer units and providing some damping so that each phase shifter is moved rapidly to the position required for the desired power output. Control of the phase-shifter output power can however be achieved more simply if the mechanical connections are changed so that the rotor of the phase shifter and the reference synchro are driven by the output shaft of integrator no. 1. Preliminary tests of this second scheme indicated that this method is feasible but that some design problems remain to be solved.

Problems of Computer Design

It becomes apparent from the preceding discussion that a computer system designed to operate in conjunction with an existing network analyzer is probably the most practical means for solving routine transient-stability problems. It is felt that the computer units of such a system should fulfill requirements of space, reliability, simplicity, cost, precision, and accuracy, and that their application should not require appreciable modifications of present analyzer design. In other words, the computer system should be an accessory to the analyzer, and not vice versa.

A complete computer installation is visualized as consisting of a group of computer units, each unit controlling a phase shifter as a regulator during steady-state operation and as a computer during transient studies. A central multichannel recording device would record rotor-position data during transient-stability studies. The computer units should be sufficiently small so that they can be mounted directly on or adjacent to each phase shifter. The reliability and simplicity of the computer installation should be commensurate with that of the analyzer itself. A suitable criterion for the reliability of the computer in daily operation is that shutdowns of the order of 1 day shall occur less than 6 times a year. Moreover, the computer should be simple enough so that an operator can obtain satisfactory results following a short period of training.

The cost of the computer installation should bear a reasonable relationship to the cost of the network analyzer and to the savings which it is to effect. Based on an average percentage of the total analyzer operating time used for transient-stability studies, a reasonable figure for the cost of the complete installation might be about 20 per cent of the cost of a network analyzer. The value of the computer units as regulators for steady-state power-flow studies depends on the individual network analyzer. This feature is valuable for the MIT network analyzer because the high internal impedance of the phase shifters frequently makes their manual adjustment time consuming.

In the design of a computer system the most important performance specifications are precision and accuracy. When operating as a regulator, the computer unit should maintain the power output of its phase shifter to within 1 per cent and the voltage to within 1/2 per cent of the desired values, as measured by the master instruments at the metering desk. Such precision is normally obtained with manual adjustment of the phase shifters, where the instruments connected at the terminals of the phase shifters are calibrated from the instruments at the metering desk. The narrow tolerances are not predicated on the accuracy of corresponding power-system data on which studies are based, but on the desire to be able to reproduce part or complete sets of readings for identical terminal conditions on identical systems set up at different times. The same criterion can be applied to the results of transient-stability studies; a reproducibility of 1-to-2 per cent is desired even though the accuracy with which the solution describes the actual system phenomena is much poorer because of the simplifying assumptions made in setting up the problem.

The specification of accuracy affects the selection of two basic components of the computer unit: the power-sensing element and the integrators. The power-sensing element can take the form of an electronic wattmeter, which has not yet been devised to the required accuracy; it can be a dynamometer wattmeter, which has the required accuracy but whose indication must be read by photoelectric, capacitance-bridge, or other methods; or it can be a watt-hour meter which serves also as an integrator of the power reading. The choice of a power-sensing element with the required accuracy and reliability is one of the key problems in realizing an actual working unit. The factors governing the selection of a suitable integrator from the large class of devices for doing such work are discussed in the following.

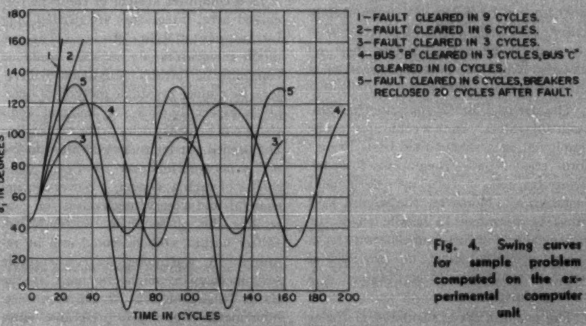


Fig. 4. Swing curves for sample problem computed on the experimental computer unit

Integrating Devices

Watt-hour meters are used as the actual computing elements of the experimental computer unit because these highly developed instruments are a simple and accurate means of integrating electric quantities. Since they cannot be loaded mechanically, a servomechanism must be used to position the output member of each integrator in accordance with the disk position. Any error-sensing device of this servomechanism which is mounted on the disk must be light in weight to keep down bearing friction and must function without imposing any other static friction torque on the meter shaft. Ordinarily, the static friction torque of the watt-hour meter bearings and register is compensated for by the light-load compensation of the meter. This compensation is unidirectional, hence, it cannot be used in computer service where the rotation of the meter disk can be both clockwise and counterclockwise. Of the available pickup devices capable of meeting these stringent loading requirements, an optical pickup was chosen for its comparative simplicity.

The choice of a watt-hour meter for performing the first integration in a computer unit is obvious since its output signal is directly proportional to the required time integral of the power difference $(P_{ik}-P_{ok})$. However, there are several devices other than a watt-hour meter which can be used to perform the second integration in the computer unit. Mechanical integrators come to mind immediately, but they are very expensive devices. The integrators of certain electronic analogue computers are another possibility, provided that their computing speed can be slowed down sufficiently for the purpose at hand. Since both mechanical and electronic integrators require follow-up servomechanisms of about the same complexity as those used with the watt-hour meter in the experimental computer, very little is gained by the use of these other integrators unless they are already available as components of a general-purpose computer.

The use of a digital computer has also been suggested. Such a computer can be employed on a shared-time basis to perform one or both integrations for each computer unit, in turn, in a permanent installation. However, the digital unit must be very fast to handle all of the phase shifters in turn and elaborate high-speed switching must be provided if accurate results are to be obtained in a reasonable time.

The use of a speed-regulated motor as an integrator was seriously considered. The output member of the first integrator would furnish the reference signal for a regulator which controls the speed of the motor used to position the phase shifter[11]; the position of the motor shaft is thus made proportional to the time integral of the output of integrator no 1. This system was not used because it was simpler to employ the watt-hour meter servomechanism design already developed for integrator no. 1.

Conclusions

The experimental computer unit described in this paper represents one of many approaches toward the development of a complete computer system for commercial service with a network analyzer. The performance of the unit is acceptable, but a considerable amount of engineering remains to be done to convert the experimental unit into a reliable and compact working model.

A complete computer installation would consist of a set of computer units, one for each phase shifter, with rotor-angle information fed from each computer unit to a central multichannel recorder. The computer units would also serve as power and voltage regulators for the phase shifters when the analyzer is used for steady-state power-flow studies. For transient-stability studies, the initial conditions would be established by the computer units acting as regulators. The units would then be converted to computer duty by setting in the machine-inertia constants and moving several levers. Starting the recorder and applying the fault to the system would start the computation. A time scale of 5-to-10 minutes of computer time for each second of system time would make recording and switching operations comparatively simple procedures.

In a computer system of the type discussed here, integrators are needed only for transient-stability studies. The use of the integrators of a separate analogue computer of suitable design may, therefore, be a more favorable method of operation, especially since the analogue computer can be utilized for other tasks when not needed for transient-stability studies on the network analyzer. For each phase shifter of such a system, a power-sensing element and a servomechanism for controlling the phase angle of the phase-shifter output voltage would still be required. A wattmeter with a suitable servomechanism might be used as the power-sensing element; however a watt-hour meter is obviously preferable, since it performs one of the required integrations as well as the power-sensing function. For steady-state power-flow studies the power-sensing element and the phase-shifter servomechanism can be combined to form a power regulator without the need for the integrators. Final evaluation of the usefulness and economic advantage of such a computer system in daily service can only be made after a network analyzer has been equipped with this system.

Appendix I. The Transient-Stability Problem

The rotors of the synchronous machines of a power system in steady state run at synchronous speed and, therefore, occupy fixed angular positions with respect to each other and to a synchronously rotating reference. A fault or other disturbance in the system results in a change in rotor speeds and angular positions. These changes are of interest, since synchronism will be lost if the angular displacement between rotors becomes too large.

The solution of a multimachine transient-stability problem thus requires the determination of the rotor angle of each synchronous machine in the system as a function of time following an initiating disturbance. This rotor angle is determined by the torque relationship governing the behavior of any rotating machine

$$\text{Inertia torque}+\text{damping torque}=\text{input torque}-\text{output torque} \quad (7)$$

Thus, for the kth machine in an n-machine system

$$J_k\frac{d^2\alpha_k}{dt^2}+\text{damping torque}=\text{input torque}-\text{output torque} \quad (8)$$

where

J_k = polar moment of inertia of the kth machine and its prime mover

α_k = rotor angle of the kth machine measured from an arbitrary fixed reference

Multiplying each term in this equation by the synchronous speed of the machine yields the following relationship[12]

$$M_k\frac{d^2\alpha_k}{dt^2}+P_{Dk}=P_{ik}-P_{ok} \quad (9)$$

where

M_k = inertia constant of the kth machine and prime mover

P_{Dk} = damping power of the kth machine and its prime mover

α_k = rotor angle of the kth machine measured from an arbitrary fixed reference

P_{ik} = mechanical power supplied to the kth machine by its prime mover

P_{ok} = synchronous power output of the kth machine measured at the air gap

Equation 9 is actually valid only if the speed of the machine is equal to its synchronous speed; however, use of this equation generally yields results of engineering accuracy in ordinary transient-stability problems where the speed does not depart materially from synchronous speed.

It is generally assumed that the damping power P_{Dk} can be neglected and that the angle

$$\alpha_k = \omega_{sk} t + \delta_k \tag{10}$$

where

ω_{sk} = synchronous speed of the machine
δ_k = phase angle of the voltage behind the transient reactance of the kth machine

With these assumptions, equation 9 can be written as

$$\frac{d^2\delta_k}{dt^2} = \frac{1}{M_k}(P_{ik} - P_{ek}) \tag{11}$$

Double integration of this equation yields equation 1.

References

1. Long Distance Power Transmission as Influenced by Excitation Systems, C. Concordia, S. B. Crary, F. J. Maginniss. *AIEE Transactions*, vol. 65, 1946, pp. 974–87.

2. Differential Analyzer Applications, F. J. Maginniss. *General Electric Review*, Schenectady, N. Y., vol. 48, May 1945, pp. 54–59.

3. The Differential Analyzer as an Aid in Power System Analysis, C. Concordia. *Paper No. 311*, CIGRE, Paris, France, 1950.

4. An Electric Analogue Method for the Direct Determination of Power System Stability Swing Curves, W. B. Boast, J. D. Rector. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1833–36.

5. Micromachines and Micronetworks. Study of the Problem of Transient Stability by the Use of Models Similar Electromechanically to Existing Machines and Systems, R. Robert. *Paper No. 338*, CIGRE, Paris, France, 1950.

6. *Discussion* by E. W. Kimbark of A New A-C Network Analyzer, H. P. Kuehni, R. G. Lorraine. *AIEE Transactions*, vol. 57, 1938, pp. 419–20.

7. A High-Frequency Simulator for the Analysis of Power Systems, S. Kaneff. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 100, part II, no. 76, Aug. 1953, pp. 405–16.

8. An Analogue Computer for Automatic Determination of System Swing Curves, D. W. C. Shen, S. Lissner. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 475–83.

9. Synchronous Machine Analogues for Use with the Network Analyzer, J. E. Van Ness. *AIEE Transactions*, vol. 73, pt. III, Oct. 1954, pp. 1054–60.

10. Analyzer Interconnections for Direct Determination of Power-System Swing Curves, G. A. Bekey, F. W. Schott. *AIEE Transactions*, vol. 73, pt. III, April 1954, pp. 238–42.

11. A Generalized Analogue Computer for Flight Simulation, A. C. Hall. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 308–20.

12. Electric Power Circuits, Theory and Applications, Vol. II, Power System Stability (book), O. G. C. Dahl. McGraw-Hill Book Company, Inc., New York, N. Y., 1938.

Discussion

Allan N. Greenwood (University of Toronto, Toronto, Ont., Canada): The authors are to be complimented on having devised what would appear to be a robust and comparatively simple analogue for a synchronous machine. A principle feature of this device is the feedback loop which causes the phase shifter to be actuated in accordance with the differential power. This, or some similar feature, is apparent in all the analogues referred to in the paper, with the possible exception of Robert's Micromachines. Essentially the device has to be "told what to do," the intelligence is conveyed by the feedback circuit to be subsequently interpreted by the integrating elements.

I should like to draw attention to what I consider a simpler device, fundamentally simpler because there is no feedback circuit. The equipment responds of its own accord to a disturbance on the analyser, and as far as possible it is arranged that its response is the same as that of an alternator confronted with a similar stimulus in a system proper. The original idea was put forward in 1938 by W. S. Stuart of the University of Leeds; he and R. F. Linsell carried out some preliminary work at that time, and demonstrated the possibilities of the device. Much later Linsell published a short note on the work.[1] More recently the project has been pursued further at Leeds by C. B. Cooper.

The device resembles a selsyn, its stator accommodates a polyphase winding, which sets up a sinusoidally distributed rotating flux when supplied with balanced polyphase currents, its rotor carries a single-phase winding which has a sinusoidal electromotive force induced in it when cut by the rotating flux of the primary. The magnitude of this electromotive force can be regulated by the primary applied voltage, its phase relative to any fixed datum will depend upon the angular position of the rotor. The rotor is connected electrically to the analyser; arranged thus it functions as a generator unit. The rotor does not rotate but takes up an equilibrium no-load position, from which it can be disturbed only by applying a torque. A torque in one direction results in a transfer of energy to the analyser, the device "generates," a torque in the opposite sense causes the device to draw power from the analyser: it "motors." The shaft torque represents the input to the machine, it can be supplied by a weight on a string around a pulley on the shaft, or more elegantly, by electromagnetic means. The power-angle curve of the model can be made to resemble quite closely the corresponding characteristic of an alternator.

In an n-machine system $(n-1)$ such generators would be connected to the network at appropriate points, the nth source would be a bus bar which would also serve as a reference. Because the generator voltages and inputs can be readily adjusted, it should not be difficult to balance such a network for a load flow study.

To simulate transient conditions with such an arrangement it only remains to endow each unit with inertia corresponding to that of the parent machine it is intended to represent. In this regard there is a certain freedom of choice, the same freedom a that enjoyed by the authors in assigning values to their constants K_1K_2 and I. It is the ratio of torque to inertia which fixes the time scale for the swings of the models when the system is disturbed. The present aim is to have them swing at the same rate as the machines proper, but this may prove difficult to achieve.

Asynchronous damping forces should be simulated with ease on such a model, and it is not inconceivable that effects of governors might also be introduced. One of the chief difficulties do date has been that of reducing bearing friction in the analogue. It is hoped that a full report on the work will be published in due course.

Reference

1. Methods of Assessing the Stability of an A.C. Power System, R. F. Linsell. *The Engineer*, London, England, vol. 192, 1951, p. 332.

James E. Van Ness (Northwestern University, Evanston, Ill.): The paper presents an interesting device for improving the usefulness of a network analyzer. The use of this type of equipment, both in transient stability problems and steady-state problems, will greatly increase the value of any network analyzer.

The watt-hour meters provide a convenient means of combining the power-sensing element and the first integrator. However, they add to both the size and cost of the unit. I have found that an electronic device such as we have used for this purpose (see reference 9 of paper) is very reliable. No mention is made of a driving amplifier for the integrator. Is the burden low enough so that it can be connected directly to the network analyzer? The authors state that the reason for using a watt-hour meter for the second integrator instead of a speed-regulated motor was that it was simpler. Might it not be worth while in a final design to use as an integrator the elements of the servomechanism that are required anyway to drive the phase-shifter? This would reduce the unit's size and cost.

While I agree that it is advantageous to reduce the time scale, I do not think it necessary to go so far as is indicated in this paper. Our system reduces the time scale by a factor of 10. At this speed, a simple relay system can easily perform switching operations, and more conditions can be evaluated in a shorter time.

In using the system for steady-state operations, we intend to leave one of the integrators in the loop instead of removing both of them, as suggested by the authors. This will have the advantage of reducing the steady-state error of the system to zero as is commonly done in servomechanism. The constants of the system can be adjusted so that the time delay introduced will be negligible.

A. Kusko and P. N. Heller: The authors wish to thank Mr. Greenwood for calling attention to further work being done, and they feel that it would be of interest to learn the final results of this investigation.

The interest shown by Professor Van Ness is appreciated. Space in a network analyzer installation is not so limited as to prohibit the use of watt-hour meters as computer elements, especially since their remarkably low cost, inherent accuracy, freedom from drift and long life give them a decided advantage over electronic watt meters in multigenerator installations. The experimental computer unit was installed on the MIT network analyzer which operated with a base current of 1 ampere. It was therefore possible to connect the potential coil of watt-hour meter no. 1 directly across the phase shifter terminals. The current coil was connected between the potential coil and the network representation on the analizer. Watt-hour meter no. 1 would, of course, be left in the circuit when the unit is used as a power regulator so that there will be an integration in the regulator loop.

The use of a watt-hour meter for the second integration was prompted by the desire to save the engineering time necessary to develop an alternate means of integration of equal accuracy. The same considerations will influence the final design since development costs form a major portion of the price of a commercial installation.

It was felt that the slow computing speed would not be particularly objectionable since the preparations for a transient stability run would still require about as much time as the actual computation. A slow computing speed improves the accuracy of the computer units and makes it possible to avoid the inaccuracies and complications of high-speed recording.

A Circuit for Measuring the Resistance of Energized A-C Windings

R. E. SEELY
MEMBER AIEE

A SIMPLE means of measuring the electrical resistance of a-c energized windings has long been desired inasmuch as it would permit the determination of winding temperatures while the windings are operating under actual load conditions. The circuit as outlined in this paper makes such measurement possible.

Discussion

Present Methods

The present procedures for determining windings or coil temperatures are either thermometers or thermocouples or to calculate of the winding temperature from winding resistance values obtained at room temperature and after shutdown. Each of these procedures has certain inherent disadvantages.

The inherent disadvantages of thermometers for measuring winding temperatures are:

1. The thermometer tends to measure only the temperature of the spot where it is located and, in general, can only be used in an attempt to measure surface temperatures of winding.
2. There is frequently a considerable difference between the temperature as measured by a thermometer and the actual surface temperature of a winding. This difference between actual surface temperature and the thermometer reading can vary considerably, depending on the manner in which the thermometer is applied to the winding.
3. Because of the poor thermal contact between a thermometer and a winding, there can be a large temperature lag between them when the winding temperature is changing rapidly.
4. The thermometer frequently cannot be applied to the windings of a small apparatus without modifying the enclosure of the apparatus.

The disadvantages of a thermocouple for measuring winding temperatures are:

1. The thermocouple measures only spot temperatures.
2. There can be a wide variation in the thermocouple temperatures resulting from the manner in which thermocouples are applied to the windings.
3. The use of different types of thermocouples and of different size thermocouple wire can also result in a considerable difference in the measured temperature.
4. There can be temperature lag between the winding temperature and the thermocouple temperature when winding temperatures are changing rapidly.
5. It is frequently necessary to disassemble the apparatus under test to install thermocouples.
6. With hermetically sealed apparatus, it is necessary to build special test units to bring out the thermocouples.

The method of determining winding temperature by measuring winding resistance after shutdown also has the following inherent disadvantages:

1. Accurate winding resistance measurements cannot be made because of the changing resistance.
2. Since resistance readings cannot be made immediately after shut down, it is general practice to obtain a number of resistance readings after various time intervals and to plot these values back to zero time on logarithmic graph paper. This procedure is time-consuming and somewhat inaccurate since the cooling curve is not entirely logarithmic.
3. Winding temperatures cannot be followed under changing load conditions by this method.
4. Equipment such as refrigeration equipment cannot be frequently shut down during tests without disrupting the entire test.

New Resistance Method

It will be realized that, if attempts were made to measure the resistance of an a-c energized apparatus using a Wheatstone bridge or an ohmmeter, the alternating current voltage would damage or burn out the bridge or ohmmeter. Even if it were possible to connect the bridge directly to an a-c energized apparatus it still would be impossible to measure the resistance of this apparatus since its resistance would be paralleled by the resistances of all of the other loads on the line and, also, by the resistance of the secondary winding of the line transformer. It was, therefore, necessary to devise a circuit which would circumvent these two basic obstacles.

The circuit in Fig. 1 illustrates the manner in which this was accomplished. First, the voltage problem was surmounted by the use of a substantially equal and opposite bucking voltage as

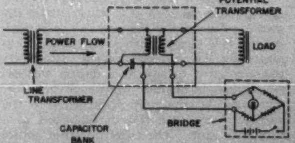


Fig. 1. Basic circuit for measuring resistance of a-c energized windings receiving power

---

Paper 55-23, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing November 12, 1954.

R. E. SEELY is with the General Electric Company, Fort Wayne, Ind.

The author wishes to acknowledge the work of P. A. Vance of the General Electric Company for designing special potential transformers and the work of Mark Van Buskirk of the P. R. Mallory Company for developing special electrolytic capacitors. He also wishes to thank the Spencer Thermostat Division of the Metals and Controls Corporation for supplying test data on a hermetic refrigeration motor.

Vo
Pt
19

supplied by a potential transformer. This allows a bridge or ohmmeter to be connected to one terminal of the apparatus on test and to one side of the secondary of the transformer since there is substantially zero voltage between these points. It was originally thought necessary to use a variable transformer to buck the voltage to zero, but it was found unnecessary since a few volts of alternating current would not adversely affect a bridge.

The necessary isolation of the windings under test so that the direct current from the bridge or ohmmeter will not flow through the other paralleling resistance in the circuit is accomplished by the use of the capacitor bank in the circuit of Fig. 1. It will be noted that the primary winding of the bucking transformer is also on the line side of the capacitor bank. As such, this permits the bridge to measure the sum of the resistances of the winding in question, the secondary of the transformer, and the leads. The actual winding resistance is then obtained by subtracting the resistances of the transformer and leads from the total resistance as obtained from the bridge. The winding temperature, of course, can then be obtained by using the equation

$$T_h = (R_h/R_c)(K+T_c) - K$$

where

R_h = hot winding resistance
R_c = cold winding resistance at the start of test
T_c = cold winding temperature as obtained from thermometer at start of test
K = constant

For temperatures in degrees centigrades, this constant is 234.5 for copper windings and 221 for aluminum windings. For temperatures in degrees Fahrenheit, it is equal to 390.1 for copper and to 365.8 for aluminum. Fig. 2 illustrates how this circuit can be used in reverse to measure the resistance of a-c energized windings which supply a-c power such as alternator windings or a transformer secondary winding.

Whereas, the circuit shown in Figs 1 and 2 illustrates the fundamentals of this circuits operation, it will be realized that further circuit refinements were necessary for successful use in a test department. The circuit shown in Fig. 3 incorporates refinements such that it can be successfully used to measure the winding resistance of either single-phase or polyphase equipment. To measure the resistance of energized polyphase equipment, it is necessary to insert another capacitor bank in the third line so that the direct current of the bridge cannot flow back through the line transformer. With grounded neutral polyphase equipment, it is necessary to go a step further and to have a capacitor bank in all three supply lines to isolate the flow of bridge current. When used with wye-connected polyphase equipment, the resistance of 2-phase windings in series is measured. With delta-connected equipment, the resistance of a phase winding paralleled by the other 2-phase winding in series is measured.

It will be noted that a 4-pole double-throw switch has been incorporated into the circuit, as shown in Fig. 3. This switch has a number of purposes. First, it serves to connect the potential transformer into the circuit when the switch is thrown to its READ or ON position. This is desirable so that the potential transformer is only energized when it is necessary to take a resistance reading. In this manner, there is very little heating of the transformer which would change the resistance of the transformer secondary winding. It should again be noted that the purpose of the potential transformer is only to supply a bucking voltage and not to supply power. As such, the temperature rise would normally be small even if it were energized continuously. In the OFF or CHECK position, as shown this switch shunts the capacitor banks *C1* and *C3* and connects the secondary of the potential transformer directly to the bridge terminals. Normally, the impedances of the capacitor banks are such that there is a few volt drop across them at the full-load capacity of these circuits. It is desirable to shunt the capacitor banks *C1* and *C3* such that there will be no voltage drop across them when the circuit is OFF. This allows voltmeters and wattmeters to be connected on the line side of the circuit such that their readings will be correct when the circuit is in the OFF or CHECK position. These meters can be connected to the load if provisions are made to disconnect them while a resistance reading is being made. Connecting the secondary of the transformer to the bridge terminals on the CHECK or OFF position permits its secondary resistance to be checked periodically for any slight change of the tare resistance. It will be noted that two blades of the switch to the secondary of the transformer are connected in parallel. This was found to be desirable to obtain a low and stable resistance through the switch.

The circuit, as shown in Fig. 3, is intended to be used to make periodic resistance readings during a heat run, but it also can be used to make continuous readings should they be desired. To do this, it is necessary that a circuit be energized for a few hours at the desired test voltage with the switch in the READ position and with no load being applied to the load terminals. This permits the potential transformer to heat and allows its tare resistance to become stabilized, as can be determined by occasionally switching to the CHECK position.

Another refinement of the circuit in Fig. 3 is the addition of a capacitor bank *C2* directly across the bridge terminals of the circuit. The use of this additional capacitor bank was found to be desirable to minimize the effect of line surges on the galvanometer of the bridge. Without this additional capacitor bank, it occasionally was found that the switching surges on the line or a pulsating motor load would cause the galvanometer to be unstable. To minimize galvanometer unstability caused by

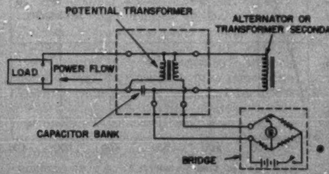


Fig. 2 (left). Basic circuit for measuring resistance of a-c energized windings producing power

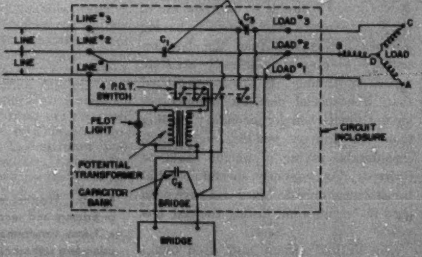


Fig. 3 (right). Refined test circuit for measuring resistance of a-c energized single-phase or polyphase windings

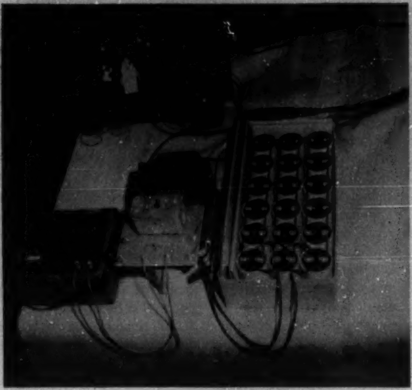


Fig. 4 (above). The original resistance measuring equipment being used to determine winding temperature of a single-phase induction motor

Fig. 5 (right). The latest circuit enclosure for measuring resistance of a-c energized windings. This circuit is rated at 20 amperes, 250 volts, 60 cycles, single- or polyphase, and 0.2 to 100 ohms

Vol. Pt. 1 1955

these surges, it has also been found desirable to operate this circuit from a "stiff" line or, better yet, to operate from a separate voltage regulating transformer or separate alternator, if available. The addition of a pilot light to the circuit is desirable so that the circuit will not be energized any longer than necessary. In this manner, heating of the transformer is very slight and, hence, the change in the transformer resistance is negligible.

In the original circuits which were built about 7 years ago, standard 1- and $1^1/_2$-kva power transformers were used because they were readily available. They were connected for 230/230-volt operation, and as such the circuits could be used for voltages up to about 250 volts. Transformers of this size were employed so that the tare resistance of their secondary winding (about 0.3 ohm) would not be greater than the resistance of the induction motor windings to be measured. It was found desirable to limit the use of these test circuits to an apparatus having a resistance greater than the tare resistance to obtain sufficiently accurate resistance reading. It was soon realized that, if a special potential transformer having a much heavier secondary winding than the primary winding were used, considerable saving in weight could be realized. Consequently, special potential transformers were designed and have been used in circuits built during the last few years.

The original circuit also used banks of standard 110-volt a-c electrolytic motor-starting capacitors, rated 850 to 900 microfarads, since they were readily available. Here again, it was realized that these capacitors were many times larger than needed, since even under full loads the voltage across them was only 4 or 5 volts. Subsequently, electrolytic capacitors rated 15 volts a-c and having a capacity of 4,000 to 9,000 microfarads were obtained. As such, it has been possible to reduce the size of the capacitor banks by a ratio of about 10 to 1. At this point it should be mentioned that the internal leakage resistance of these capacitors, in effect, parallels the resistance being measured, and establishes the maximum resistance which can be measured with sufficient accuracy. It is for this reason that a maximum resistance rating is applied to a given circuit.

Fig. 4 shows one of the original circuits being used to determine the temperature of a motor winding and the various circuit components. Fig. 5 shows the latest type of measuring circuit box which uses the circuit shown in Fig. 3 and the special transformer and capacitors described in the foregoing. This circuit box has been rated at 20 amperes, 250 volts, 60 cycles and 0.2 to 100 ohms resistance. It weighs 116 pounds, most of it being the transformer. Although this circuit box has been rated at 60 cycles, it has been used at frequencies from 25 to 430 cycles. When used at other frequencies, the volt-

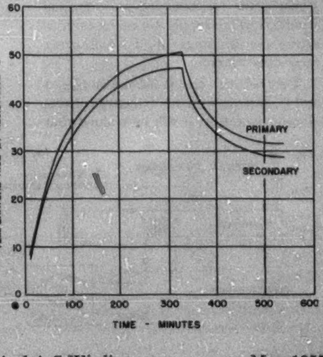


Fig. 6. Heating and cooling curves of primary and secondary windings of a 1-kva transformer as obtained by using two resistance measuring circuits

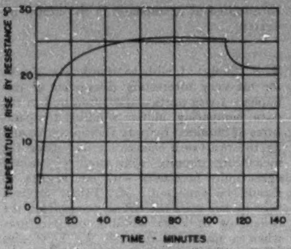
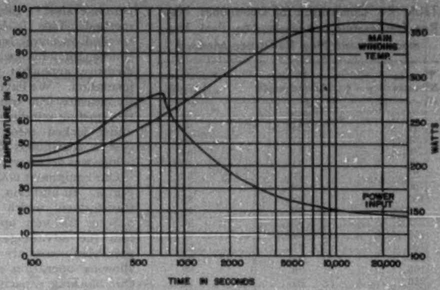


Fig. 7 (left). Temperature rise curve of windings of a 1.5-kva 3-phase 115-volt 430-cycle motor alternator which was first loaded to full load and then to 50-per-cent load

Fig. 8 (right). Curves of winding temperature and power input of a 1/6-horsepower hermetic refrigerator motor



age and current rating of the circuit is varied by the ratio of the odd frequency to the rated 60-cycle frequency.

Although, the measuring circuit was originally developed to determine the temperature of motor windings by the resistance method, it also has been used to determine the winding temperature of alternators and transformer windings. Fig. 6 shows the heating and cooling curves of the primary and secondary windings of a 1-kva transformer, as obtained by using measuring circuits in both the primary and secondary windings. In this test, the transformer was first loaded to full and then reduced to 50-per-cent load after $5^{1}/_{2}$ hours. Fig. 7 shows a temperature rise curve of the winding of a 1.5-kva 3-phase 115-volt 430-cycle motor alternator which was first loaded to a full load and then to 50 per cent load.

Fig. 8 illustrates a very important use for this measuring circuit. Shown here are curves of winding temperature and power of a 1/6-horsepower hermetic refrigerator motor, as obtained when the refrigeration system reduced the refrigerator cabinet from room temperature of approximately 40 degrees centigrade to its normal operating temperatures. It will be noted that, under such a test, the power and winding temperature reach a peak and then fall off as the cabinet temperature is pulled down. Without the use of this resistance-measuring circuit it would be difficult or impossible to determine accurately the peak winding temperatures.

Conclusions

Since its conception, this circuit for measuring energized a-c winding resistance, as discussed in the foregoing, has proved to be a highly desirable piece of test equipment, and its use is continuing to expand. Because of this circuit, the method of determining the temperature rise by the resistance change method is becoming more widely accepted.

◆

Discussion

R. F. Fricke (Westinghouse Electric Corporation, Lima, Ohio): The measurement of motor temperatures under load has always been a difficult problem. Several methods of measurement are currently in use, and each method has its particular advantages and disadvantages. The thermometer method defined and used by AIEE and the National Electrical Manufacturers Association is simple and direct. However, winding temperatures measured by thermometer are necessarily low since the bulb cannot be placed in small, high-temperature cavities. Thermocouples usually permit us to associating the measuring device more closely with the source of heat. As a result, this method is quick-acting, and temperatures measured by thermocouple more nearly approach the actual temperature of the element measured. However, test results with thermocouples depend upon several variables, one of which is the ability to make good connections with the surface of the element to be measured. Furthermore, unless thermocouples are wound into the machine, only end-turn temperatures can be measured and the possible high-slot temperatures remain unknown. Temperature indicated by resistance readings taken after shutdown, then extrapolated to shutdown time, gives a more accurate evaluation of over-all average winding temperature. The greatest drawback to this method is, of course, the fact that only final temperature can be determined.

The resistance method of temperature measurement as described by the author solves many of these problems and adds only few. This method enables us to measure winding resistance while a motor is running and, by comparing with the cold resistance, calculate the total winding temperature. It should be noted, furthermore, that in all single-phase motors only the main winding temperature is measured. The Wheatstone bridge current is conveniently blocked from the motor auxiliary winding by an open starting switch or by the motor-running capacitor. Thus we gain the advantage of measuring the temperature only of the hot main winding—rather than some average of both phase windings.

The resistance method of temperature measurement presented in this paper is free of some of the difficulties encountered with thermocouples. It measures over-all average winding temperature and is not dependent upon location or the connection of a temperature-measuring element. This is of particular importance for correlation purposes when temperatures are measured repetitively, for instance, by a manufacturer, his customer, and the Underwriters' Laboratories. Inherent accuracy is as good as in other methods. Perhaps the most important disadvantage of the resistance method is its inability to measure hot-spot temperature.

When we investigated the use of this novel circuit several years ago, we felt it necessary to compare this method with the thermocouple method which then was, and probably still is, the most popular method of measuring fractional horsepower motor temperatures. We conducted heating tests on several motors while simultaneously measuring temperature by both methods. Stabilized winding temperatures from some of these tests are shown in Table I. It can be seen that temperature measured by resistance compares very favorably with that measured by the hottest thermocouple. This is fortunate, of course, because peak

Table I. Comparison of Winding Temperatures as Measured by Peak Thermocouples and Resistance Method While Motor Is in Operation

Motor Type	(A) Peak Thermocouple	(B) Resistance Method	(A)/(B) Ratio
Enclosed	99.5	100	100.5
	110	110	100.0
Open 4-pole	106	107	101.1
	117.5	117	99.5
	106.5	110	103.3
	106	107	101.0
Enclosed	107	106	99.0
	111	111	100.0
Open 2-pole	49	52	106.0
	55	56	101.8
	63	61	96.8
	70	72	103.0
	84	80	95.3
	104	98	94.2
Average	92.04	91.93	100.2

Table II. Comparison of Continuous Heat Run Temperature as Measured by Peak Thermocouple and New Resistance Method

Time in Minutes	Peak Thermocouple	Resistance Method
0	25	25
15	41	50
30	47	60
45	64	68
60	71	75
75	77	82
90	82	86
105	87	89
120	90	91
135	93	93
150	94	95
165	97	97
180	97	98
195	98	99
210	99	99
225	100	100
240	100	100

thermocouple or hot-spot temperature, rather than average thermocouple temperature, is the limiting factor in motor design. We obtained the greatest difference between the two methods in open motors where the variation in air movement over the windings moved the hot spot, or caused it to vary farther from the average winding temperature.

Table II is a continuous record of a heat run, and shows that the initial rate of temperature rise is greater by the resistance method. This is logical, for the latter method measures temperature at the heat source (the wire itself) while thermocouples measure somewhat remotely on the wire insulation. We have also used this method to measure fast transient temperatures of the starting windings in split-phase motors under locked rotor conditions. For this case a continuous dial type of resistance bridge was necessary.

Our equipment is similar to that shown in Fig. 3. In addition, we use doubly wound transformers which can be switched for 115- or 230-volt operation, and thus obtain the advantage of a low resistance paralleled secondary on low voltage, while allowing operation at the higher voltage. Our blocking capacitor is made up of two 10,000-microfarad elements, one of which can be externally disconnected from the circuit.

Both capacitors are used on large low-resistance motors to reduce the unbalancing voltage drop across the capacitors. One capacitor is used on small high-resistance motors to increase the total leakage resistance of the elements. In this way the range of resistance which we are able to measure with the circuit is effectively doubled.

The author is to be commended on his interesting paper. I hope that it will serve to publicize this method so that it soon may be adopted as a standard in the motor industry.

R. E. Seely: I wish to thank Mr. Fricke for his very interesting discussion of this paper. I was particularly interested in the data comparing motor winding temperatures as obtained by peak thermocouple and by the resistance change method using the measuring circuit. The results obtained agree quite favorably with our results, as should be expected. Mr. Fricke advised that he has used the circuit to measure the resistance of start winding temperatures when a motor is stalled. We have also made a similar investigation, but wish to point out that erroneous results can be obtained if the winding resistance changes too rapidly. This error is apparently caused by the charging of the capacitor bank.

Mr. Fricke advises that in the circuits he has built, doubly wound transformers were used which can be switched for 115- or 230-volt operation. He also uses a switch for varying the amount of blocking capacitance. We have considered such additional refinements in the circuits which we built, but have not employed either because we felt that it would be desirable to keep the circuit simple to operate. Instead of varying the voltage rating of a circuit by switching, we have built individual circuits rated at 120, 240 and 480 volts.

The Propagation Mechanism of Impulse Creepage Discharges Over Oil-Immersed Surfaces

J. G. ANDERSON
MEMBER AIEE

T. W. LIAO
MEMBER AIEE

Synopsis: A fundamental study of the mechanism of propagation of impulse creepage streamers over an insulating surface is reported. The studies were made using high-speed cameras, still cameras, and special probes, to give a new picture of the creepage process. High-speed photographs show that impulse creepage in the nonuniform field is a stepped discharge process, remarkably similar to natural lightning except on a greatly reduced scale. Probe measurements display the variation of channel voltages with time and distance. The mechanism is discussed in considerable detail for both positive and negative polarities.

THE designer of electric apparatus is confronted with two types of electrical failure of solid insulation; puncture and creepage. Electrical failure by puncture has been studied extensively for many years, and its mechanism is becoming increasingly clear, at least in the case of pure solids.[1] However, knowledge of the mechanism of creepage failure is not nearly as complete. Considerable work has been done on surface discharges in gases[1,2] but a dearth of fundamental work exists on liquids. The latter is particularly true for impulse creepage over oil-immersed surfaces, a case of prime importance in oil-filled apparatus such as transformers and circuit breakers.

The few workers in this field have generally used a Lichtenberg figure or photo paper technique wherein the surface discharges move across or very close to a photo-sensitive surface.[3-5] This method has yielded some interesting results, but suffers from the fact that it does not give a record of the discharge with time. In addition, the propagation of discharges across a photo-sensitive surface may be different from the propagation across surfaces of practical insulation.

To study the propagation of these discharges over insulating surfaces, an investigation was undertaken with the use of high-speed and still cameras to record the streamer progression, and using a special probe technique to measure the rise and fall of potential in the creepage channel as it progressed. The results have given a new and clearer idea of impulse creepage in oil and hold promise of revealing new facts about the breakdown of liquids in general.

In studying the fundamental process of impulse creepage, a nonuniform field electrode configuration was used. Such a field was chosen for several reasons:

1. Experience has shown that in such a field, the velocity of the creepage streamers is slow enough to permit detailed measurements with available equipment.

Paper 55-36, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing November 16, 1954.

J. G. Anderson and T. W. Liao are with the General Electric Company, Pittsfield, Mass.

The authors wish to acknowledge the direction and encouragement given by J. H. Hagenguth and T. R. Walters, and the active support and advice of their many associates, particularly R. B. Kaufman, F. I. Winslow, J. E. Eulian, A. H. Sharbaugh, and J. C. Devins.

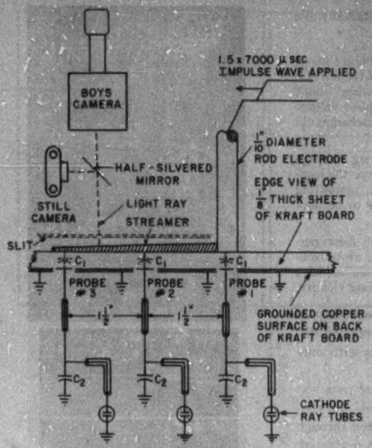

BOYS CAMERA
1.5 x 7000 µ SEC IMPULSE WAVE APPLIED
1/10" DIAMETER ROD ELECTRODE
HALF - SILVERED MIRROR
STILL CAMERA
LIGHT RAY
STREAMER
EDGE VIEW OF 1/8" THICK SHEET OF KRAFT BOARD
SLIT
C1
PROBE #3
PROBE #2
PROBE #1
1 1/2"
GROUNDED COPPER SURFACE ON BACK OF KRAFT BOARD
C2
CATHODE RAY TUBES

Fig. 1 (left). Sketch of instrumentation techniques for simultaneous streamer channel potential measurements and photography

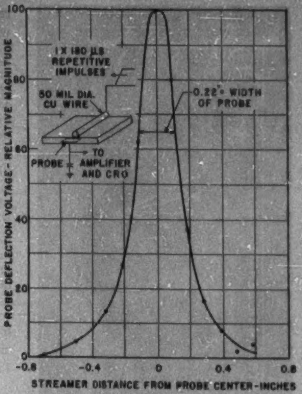


Fig. 3 (right). Probe calibration

2. The discharge in a nonuniform field oil gap has been shown in a previous paper[4] to consist of stepped streamers similar to natural lightning. For comparison purposes it was found desirable to use a nonuniform field for this investigation also.

3. Nonuniform field creepage is a common phenomenon in the breakdown of oil-immersed insulating structures.

Instrumentation

A sketch of the basic technique used in these studies is shown in Fig. 1 and a photograph of the instrumentation in Fig. 2. A sheet of Kraft board insulation 14 by 14 by 1/8 inches was mounted vertically in a large lucite tank filled with commercial quality *10-C* oil. On one side of the insulation a flat-ended steel rod electrode 1/10 inch in diameter was pressed against the surface. This rod was connected to an impulse generator, and produced impulse streamers which moved out over the insulation surface. The other side of the insulation was covered with a grounded film of copper 4 mils thick, deposited directly on the material. On this copper surface three 0.2-inch square areas, $1^1/_2$ inches apart, were isolated from the rest of the grounded film by a 1/32-inch insulated boundary. The copper surface was directly connected to ground, but the isolated areas or probes were connected to ground through capacitors across which the plates of the cathode ray tubes were connected. Fig. 1 shows that the equivalent capacitance C_1 from the insulation surface to the probe and the capacitance C_2 from probe to ground formed a simple capacitance divider. C_2 was made about 1,000 times C_1, so that for every kilovolt of streamer potential on the barrier surface, 1 volt appeared across C_2. This voltage across C_2 was fed to the vertical deflecting plates of the cathode-ray tubes for display. The sweeps of the cathode-ray tubes were initiated by the impulse wave applied to the rod electrode. Thus a record of the approximate streamer potentials with time could be obtained. The low probe potential had negligible effect on the impulse creepage process.

Figs. 1 and 2 also show a high-speed Boys camera and a still (Leica) camera which were mounted outside the transparent lucite tank to photograph the discharges as they moved over the insulation surface. The high-speed Boys camera has been described in previous papers,[6,7] and was used to obtain a con-

Fig. 2 (left). Experimental arrangement

A—Herkolite bushing
B—High-speed Boys camera
C—Still (Leica) camera
D—Rod electrode
E—Insulation
F—Lucite tank
G—Oil level

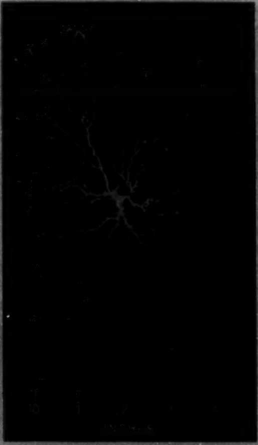


Fig. 4 (right). Still photograph (without slit) of 120-kv negative-polarity creepage streamers on Kraft board surface

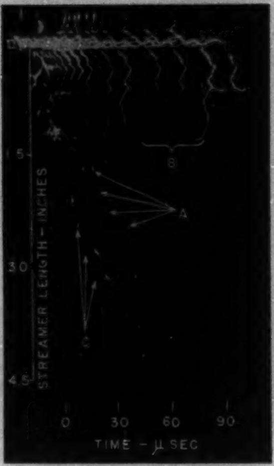


Fig. 5. High-speed Boys camera photograph (through slit) of single 120-kv negative-polarity creepage streamer

A—Stepped streamer
B—Streamer passing beyond edge of slit
C—Glow at tips

tinuous time sequence of the progression of the impulse creepage discharge. In principle it consisted of a circular strip of film moving at high speed past a shutterless *f 2.3* lens. The film speed of 1.6 to 2.0 mils per microsecond allowed resolution of events 1 microsecond apart in time. A shutter was not necessary since all tests were made in darkness, and the discharge itself supplied the light recorded on the film.

The Leica camera was mounted at right angles to the Boys camera, and the discharge light fed into both cameras by use of a half-silvered mirror. This, of course, reduced the light to the Boys camera, and therefore in many cases the mirror was removed.

Fig. 1 also indicates that a 5/16-inch wide slit in line with the capacitance probes was arranged between the insulation and the camera to mask out the light from all discharges other than those passing over or near the probes. This greatly simplified the interpretation of the mechanism of progression, but necessitated photographing a large number of discharges before one would pass under the slit for any considerable distance.

During all tests, an impulse wave with a 1.5-microsecond front and at least a 7,000-microsecond tail was used. The reasons for using such a long-tailed wave were threefold:

1. It approximated a step-function wave, making a theoretical analysis of the impulse creepage propagation mechanism much easier because variations in applied voltage during the propagation of the discharges did not have to be considered.
2. With a long-tailed wave, long creepage discharges could be obtained over the insulation surface at a low voltage, permitting much thinner insulation sheets to be used without puncture.
3. The use of thinner insulation sheets permitted smaller probes and better probe resolution, so that the probe became sensitive only to discharges in its immediate vicinity.

With such a long-tailed wave it was found possible to obtain creepage discharges as long as five inches with only 120-kv impulse voltage.

The immersant used in these tests was good quality commercial *10-C* oil continuously circulated and filtered through a felt filter. The oil was not deaerated, since this would have greatly complicated the tests, but the insulation sheets themselves (after the probes were constructed) were treated with deaerated oil under vacuum and transferred to the testing tank under oil when the tests were started. During these experiments, the insulation surface was wiped free of residual charge with a grounded copper wire before the initiation of each creepage streamer discharge to prevent any accumulative charge effects from influencing the results.

Probe Calibration

To estimate the creepage streamer potentials, the capacitance divider C_1-C_2 was calibrated by placing a flat electrode on the insulation surface opposite the probe. A known voltage was applied to this electrode and the resulting oscillograph deflection measured. Deflection sensitivities of about 0.005 inch per kilovolt were obtained. Of course such a calibration was more accurate for large streamers having diameters approaching the 220-mil probe width. For streamers of smaller diameter an error was introduced because of the smaller streamer capacitance C_1, becoming about 50 per cent for a 50-mil diameter streamer. Thus the streamer potentials were always at least as high as those indicated by the probes and in many cases considerably higher. However, by observing the streamer channel diameter photographically, and allowing for effects of adjacent channels, branching, and other factors which tended to increase C_1 to a higher value, a better estimate of streamer potentials was possible.

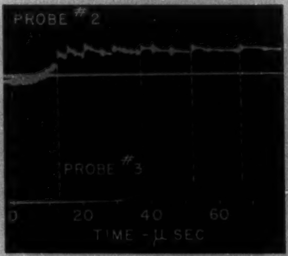


Fig. 6. Oscillograms of 120-kv negative-polarity creepage channel potentials. Crest voltage at probe 2 = 120 kv; crest voltage at probe 3 = 85 kv attributable to streamer not passing directly over probe

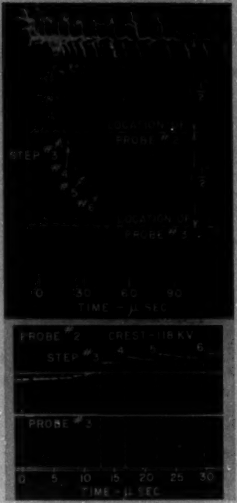


Fig. 7. Boys camera photograph and probe oscillograms of 120-kv negative-polarity creepage streamer

The probe resolution with regard to streamer location was determined by a wire 50 mils in diameter placed at various positions over the probe to simulate an impulse streamer. This simulated streamer was impulsed with 1 x 180-microsecond repetitive impulses of 1,000-volt amplitudes. The resulting probe voltages were measured by means of an amplifier and oscillograph at various wire positions parallel to the slit. Fig. 3 shows the resulting voltages measured. It may be seen that the probe sensitivity

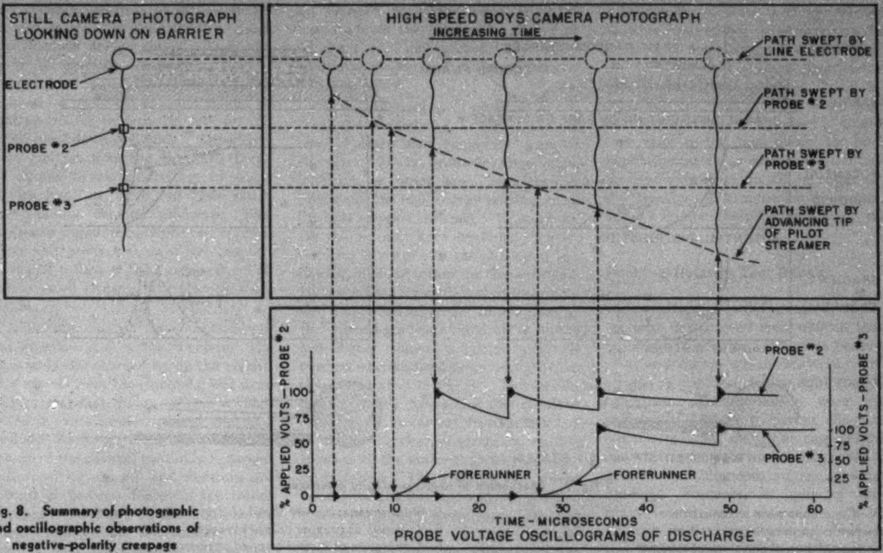


Fig. 8. Summary of photographic and oscillographic observations of negative-polarity creepage

drops rapidly as the wire is displaced from the center position, becoming 65 per cent at the probe edges and 25 per cent at a distance of one probe width away from center.

Negative-Polarity Test Results

When a negative-polarity impulse wave was applied to the rod electrode, impulse creepage streamers moved outward over the insulation surface in all directions from the electrode end. A typical still photograph of these streamers, taken without the mask and slit, is shown in Fig. 4. Here one can observe both the shrinkage of the creepage streamer channels as they moved out over the insulation surface, and the extensive branching that occurred.

When a single discharge was photographed through a slit by the Boys camera, its propagation over the insulation surface could be resolved with time. A typical result is shown in Fig. 5. Here, increasing time is to the right and the lengths and times of occurrence of the streamers are as shown. This photograph reveals that a single discharge (as seen by the eye or by a still camera) is actually composed of many bright streamers, called stepped streamers, all moving down the same channel with ever-increasing length. Each streamer exists for only 1 microsecond or less, and the interval between streamers generally increases with time from 1 microsecond or less at the beginning to 10 microseconds or more toward the end of the process. Furthermore, when one streamer branches, subsequent streamers usually branch at the same point, and most of the streamers appear to have bright spots at their tips. This type of propagation is commonly referred to as the "stepped" discharge, and has been found extensively in natural lightning,[7,8] in point-to-plane gaps in air[9] and recently in oil.[6,10]

A typical set of negative polarity discharge oscillograms is shown in Fig. 6. Probe 2 measured the channel potential at a distance of 1.5 inches from the origin at the rod electrode, and probe 3 measured the potential at a distance of 3.0 inches from the same point. On these oscillograms a small displacement of the zero line can be observed. This displacement is attributable to a small residual charge collection on capacitor C_2 during the tests, but it in no way modifies the probe frequency response to the potential variations of interest. It merely shifts the d-c response, and points out the desirability of having resistance compensation across C_2 in future tests.

In Fig. 7, a set of simultaneous high speed Boys camera photographs and oscillograms are shown. Comparison of the times of occurrence of the stepped streamers and sudden voltage increases on the oscillograms leads to the conclusion that the two are simultaneous, i.e., at the same time that the channel becomes a good conductor it emits light, resulting in a stepped streamer on the photograph. In this particular set of oscillograms, the voltage variations at probe 2 are shown in considerable detail. However, by the end of the sweep, the discharge is just beginning to arrive at probe 3, hence the latter shows no appreciable voltage variations in this short time.

Postulated Mechanism of Negative Impulse Creepage

Fig. 8 summarizes the results of photographic and oscillographic observations of a single streamer. A single creepage discharge at the top left of the sketch is one which might be observed by eye or by still camera. This discharge is shown by the Boys camera image at the top right to be a series of stepped streamers of ever-increasing length passing down the channel. Their velocity is extremely high and cannot be resolved by the camera. However, the over-all propagation velocity is much slower, because of the intervals of several microseconds between streamers, and 0.1 inch per microsecond is an average value.

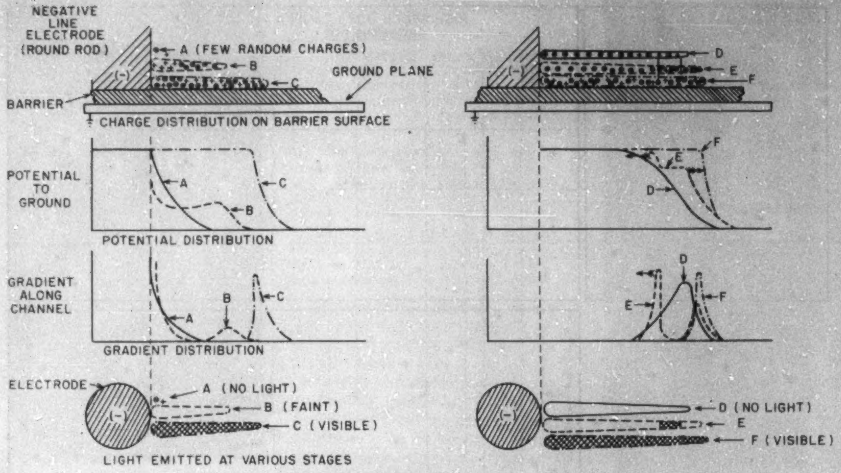


Fig. 9. Mechanism of negative-polarity creepage

A—Conditions prior to ionization
B—First pilot streamer propagation
C—Completion of first stepped streamer
D—Pilot streamer moving ahead of dying stepped streamer
E—Initiation of streamer channel breakdown
F—Completion of second stepped streamer

Vol.
Pt. 1
1955

The probe oscillograms in Fig. 8 reveal several new facts about the creepage mechanism. For negative polarity the potential above the probe is essentially zero until the creepage process begins to arrive at that point. Then, in spite of the fact that the Boys camera has shown that the streamers propagate in a series of discrete steps, the potential above the probe first starts increasing in a gradual and practically uniform manner. This observation becomes of great importance when it is recalled that the stepped leaders of lightning are believed to be preceded by a faintly luminous, continuous discharge[8] (called the pilot streamer). It has been found difficult to explain the stepping process in lightning without the pilot streamer although this has never been photographed. The Boys camera shows that in oil, the same type of stepping discharge occurs, and the probe potential measurements indicate a continuous charge propagation ahead of the steps. Thus it appears to be a possible confirmation of the pilot streamer theory of the stepped discharge and further clarifies the creepage mechanism. Therefore the propagation of the apparent pilot streamer is sketched in Fig. 8 in dashed lines, since it is not luminous enough to be recorded by the Boys camera.

After the gradual build-up of pilot streamer potential, the channel potential suddenly jumps to practically full-line voltage, indicating that the channel has suddenly turned into a good conductor. The channel emits light momentarily at this instant, forming a stepped streamer and this light is recorded by the Boys camera. However, the light emission stops almost immediately, and the channel voltage over the probe begins a slow decay. (By this time the tip of the advancing pilot streamer has moved considerably beyond the point of probe measurement.) The channel potential decays for several microseconds, then the channel suddenly restrikes, a new stepped leader of longer length is formed, the channel potential again increases to approximately the applied voltage, and the process repeats. In general, the interval between steps seems to increase markedly as time increases, although the over-all propagation velocity appears to vary only from 0.1 inch per microsecond to 0.05 inch per microsecond.

Eventually, the pilot streamer-stepped streamer mechanism arrives at the next probe, and the same type of voltage variation is again recorded. Here, however, the rise of pilot streamer voltage appears more gradual and persists for a longer time, corresponding to the longer interval between stepped streamers. Finally, a stage will be reached where no more stepped streamers occur. The channel charge then slowly leaks to ground, in a period of several minutes, through the normal insulation leakage paths.

Based on the foregoing data, a step-by-step description of the discharge propagation can now be postulated. This decription is shown in Fig. 9, where, at *A*, the potential distribution and gradients along the insulating surface are sketched before any ionization or corona occurs. The gradient is highest at the electrode edge, and rapidly decreases with distance along the surface. (It should be noted that this entire process is taking place at initial gradients about three times the critical corona starting gradient.)

At *B*, a stray electron in the high gradient region has struck a neutral oil molecule knocking off an additional electron. Each of these electrons strike other molecules, knocking off still more electrons. The electrons, being more mobile, are forced away from the electrode toward the tip of the streamer by the high field, leaving the immobile positive ions behind.

The electrons near the tip produce a sufficient field to enable the pilot streamer to continue its progression. The positive ions enhance the field at the electrode. As the pilot progresses further, more positive ions are produced. Thus the electrode gradient is continuingly increased until it becomes so high that it results in a high field emission similar to that postulated by Von Hippel[11] for the initial breakdown of liquids. These electrons emitted from the electrode charge the channel to near line voltage, causing a sudden jump in probe potentials and a reduction of channel gradients. This facilitates rapid recombination of electrons and positive ions in the channel, giving off a flash of light and heat. The first stepped streamer therefore is completed.

After the charges have recombined, the situation at D of Fig. 9 exists. The channel is still charged up by the surplus of electrons from the electrode, but is no longer emitting light. Now, a high gradient exists at its tip, and the pilot streamer continues from this point. As the tip of the channel continues to propagate, positive charges and electrons are formed as before. Electrons are forced outward toward the tip and the heavier positive charges are left behind. However, a new complication now arises. The channel gradients near the tip begin to pull electrons from the old stepped streamer channel. As these electrons move forward to fill the advancing tip of the pilot streamer, the electron density in the old channel starts dropping. This results in the gradual fall in channel potential shown in the oscillograms after a stepped streamer has occurred. Accordingly, gradients must now be increasing along the channel as shown at D in Fig. 9.

At this point, it is interesting to note that the charge concentration at the tip of the streamer has been postulated as primarily electronic. One may also conclude that, after Meek,[2] the streamer tip will continue to progress as long as the tip gradient exceeds the minimum breakdown strength of the dielectric. For a very small-area electrode such as the streamer tip might represent, the negative polarity impulse breakdown strength of oil might approximate 1,500 volts per mil. Electron mobilities in oil are not well known but a value of 1 centimeter per second per volt per centimeter might be representative, when one extrapolates from the known electron mobilities in other liquids. When the probable tip gradient is multiplied by this extrapolated mobility (see equation 5), a tip velocity is arrived at of about 0.59 centimeter per microsecond, or 0.23 inch per microsecond. This calculated value of tip velocity is of the same order of magnitude as the measured velocity of 0.1 inch per microsecond, and gives additional assurance that the mechanism is as described.

During the build-up of gradients in the channel described at D in Fig. 9, the old channel must have somewhat changed character. It is probable that some hydrogen or other gases have formed because of the high energy released from the first stepped streamer. In addition, the channel may have expanded slightly, resulting in more of a gas phase and less of a liquid phase within the channel itself. Its breakdown strength is therefore probably somewhat lowered. In any case, as the channel gradients build up to higher and higher values, a critical point is reached wherein breakdown occurs in the channel as at E in Fig. 9. This breakdown probably occurs near the tip of the old stepped streamer channel and, by distorting the channel gradients, rapidly grows until the entire channel turns into an avalanche. Electrons are then swept toward the channel tip, and the positive charge left behind again causes electron emission from the electrode surface. This electron emission charges up the channel to line potential for a second time as at F in Fig. 9. The channel gradients fall to zero, positive and negative charges recombine, and light is emitted, forming the second stepped streamer.

This alternate propagation of pilot streamer and stepped streamer must then continue as before, but cannot continue indefinitely. Some electrons attach to impurities in the oil or on the insulating surface, becoming negative ions which are relatively immobile and cannot feed the advancing tip. Others are shunted to other channels by branching. Still others lose their energies by molecular excitation and other processes. The net result is that the supply of electrons to the tip of the pilot streamer diminishes as the streamer advances. This diminution in electron charge at the streamer tip must cause the tip to shrink in diameter, if the tip gradient is to remain sufficient to keep the oil breaking down. As the streamer shrinks in diameter, it takes a longer time to pump an equal amount of total charge from the old stepped streamer channel into the shrunken pilot streamer ahead of it, and consequently a longer time for the channel gradients to build up to a value sufficient to cause a new stepped streamer to form. This theory of the process is advanced as an explanation of the increase in interval between stepped streamers as the number of steps increases.

Eventually a state will be reached wherein the falling gradients in the streamer do not feed enough electrons into the tip to maintain a critical tip gradient. The streamer channel then gradually decays, and the limit of impulse creepage has been reached. Of course, if at any time prior to this final decay, the pilot approaches near a grounded metal surface the gradients at its tip will be enhanced, the streamer will turn toward the surface, and a complete creepage breakdown will ensue.

Positive-Polarity Test Results

Acquisition of valid positive-polarity creepage data proved more difficult than for negative polarity. This was because positive-polarity streamers were more erratic in direction, length, and channel brightness. The streamers were not sufficiently straight to permit adequate slit photography, and in no case did the same streamer pass over both probes to give two simultaneous streamer voltage readings. However, in spite of this, many records were obtained which indicated that positive-polarity creepage has considerably different characteristics from those of negative-polarity creepage.

A typical set of positive-polarity channel oscillograms is shown in Fig. 10. The upper oscillogram is from probe 2 and the lower from probe 3. In this case, probes 2 and 3 were not measuring the same streamer as that already discussed. However, by comparison with Fig. 6 it can be seen that positive-polarity creepage streamer potentials build up in a different manner than those of negative polarity. The sudden jumps in potential are still present, but the increase is essen-

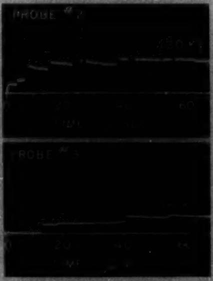


Fig. 10. Typical oscillograms of 120-kv positive-polarity creepage channel potentials. Crest voltage at probe 2=120 kv; crest voltage at probe 3=96 kv

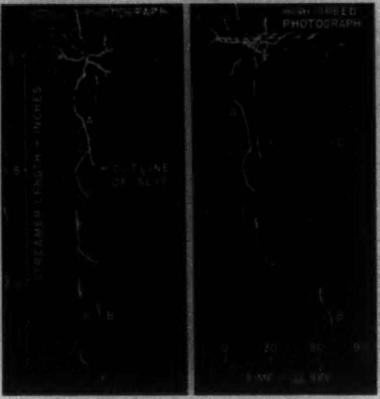


Fig. 11 (left). Still and Boys camera photographs of a positive-polarity discharge passing behind the slit

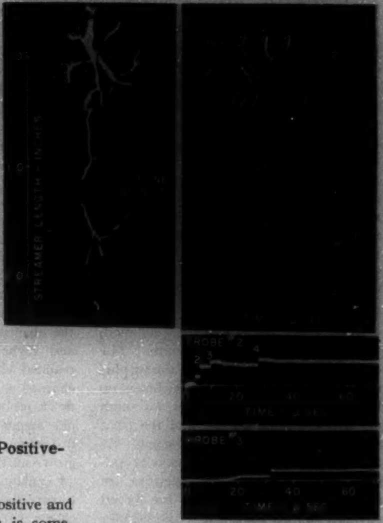


Fig. 12 (right). Still and Boys camera photographs and probe oscillograms of 120-kv positive-polarity creepage streamer

tially steplike, and the gradually increasing forerunner voltage (pilot streamer) is missing. Also the decay of potential after each increase is generally smaller than that for negative polarity.

In Fig. 11 are shown a still photograph and a high-speed photograph of a positive-polarity discharge passing down the slit. In the still photograph on the left, the edges of the slit appear in dotted lines and several streamers can be seen behind the slit. These streamers are shown by the high-speed photograph on the right to consist of many discharges, separated in time, but frequently passing down the same channel. Their velocities are erratic, some streamers, such as streamer *A* reaching distances of 1.5 inches or more in a few microseconds, whereas streamer *B* requires almost 50 microseconds to reach a distance of 4 inches. In general, the average propagation is faster in the region near the electrode. However, judging from many photographs and oscillograms, an average velocity of 0.1 inch per microsecond is estimated. Branching is also erratic, that at point *C* being a typical example. Occasionally, discharges have been observed to ignore a section of a previously formed channel and strike out in a new direction. This type of behavior appears to be because a breakdown takes place from the surrounding oil to the channel rather than from the channel tip outward.

Fig. 12 shows a simultaneous set of still and high-speed photographs and streamer channel voltage oscillograms. Here again, probes 2 and 3 may not be measuring potential variations in the same streamer channel and hence cannot be compared with confidence.

Postulated Mechanism of Positive-Polarity Creepage

The general picture of the positive and negative streamer mechanism is somewhat the same. However, there are several distinct differences because of opposite electronic movements for the two polarities. First, the channel of a positive stepped streamer does not jump up near line potential on the first discharge but builds up in steps after each stepped streamer. This is apparent because the measured channel potential is proportional to the net positive charge (above the probe), which does not come from the electrode emission as for the negative polarity but depends on the degree of ionization; consequently more charges are available after successive ionizations. Also, there is no pilot streamer ahead of the stepped streamers. Finally, the channels are more crooked and the branchings more random than those of negative polarity. This can be explained by a step-by-step description of the streamer propagation as shown in Fig. 13.

At *A* in Fig. 13, an electron is shown being attracted to the high-gradient region near the electrode, initiating an electron avalanche. Electrons are swept into the electrode, leaving positive charges behind as shown at *B* in Fig. 13. These charges effectively extend the electrode potential into the oil, thus shifting the maximum gradient away from the electrode. Concurrently photon electrons are produced by virtue of recombination in the channel, which also produces visible light. The photon electrons are again attracted toward the electrode, causing an extension of the channel. The process then repeats. However, as the channel progresses farther, the maximum gradient decreases in magnitude until it falls below some critical value where further ionization is not possible. This is illustrated at *C* in Fig. 13. All this constitutes a stepped streamer which occurs within 1 microsecond. Such a fast process is not resolvable by the Boys camera. Nevertheless the postulation should be considered reasonable when it is compared with the theory of positive streamer propagation in air,[3] because, up to this point it is difficult to see any appreciable difference, except for magnitudes, between the two mechanisms.

Since there is no pilot streamer to pave the way ahead of the stepped streamer, the stepped advance must depend on the space charge build-up along the channel. Upon examination of the probe oscillograms it is noticed that the probe potential may rise or fall with time between two successive stepped streamers. This indicates migration of charges in the channel away from or toward the probe. It is conceivable that the migration is not uniform along the channel because the arrested streamer channel is not uniformly ionized as seen from the high-speed photographs. Thus it is possible to build up more than one space charge center along the channel as illustrated at *D* in Fig. 13. For clarity, only the net positive charges

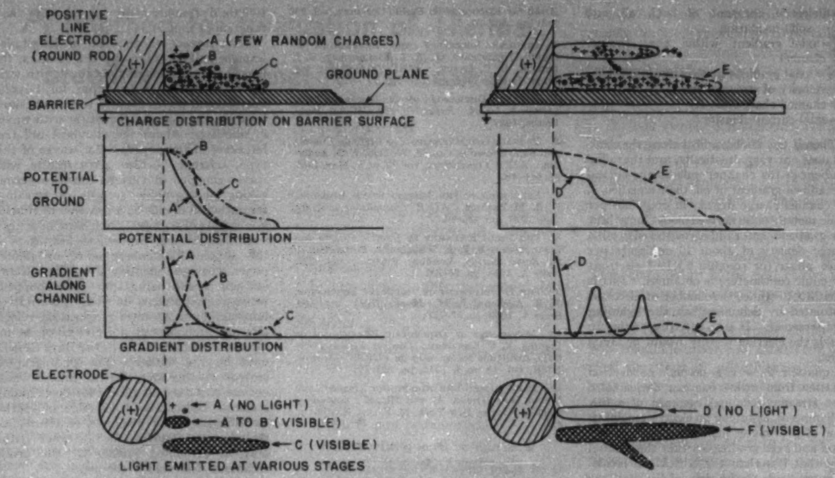

POSITIVE LINE ELECTRODE (ROUND ROD)
(+)
A (FEW RANDOM CHARGES)
B
C
GROUND PLANE
BARRIER
CHARGE DISTRIBUTION ON BARRIER SURFACE
POTENTIAL TO GROUND
POTENTIAL DISTRIBUTION
GRADIENT ALONG CHANNEL
GRADIENT DISTRIBUTION
ELECTRODE
A (NO LIGHT)
A TO B (VISIBLE)
C (VISIBLE)
LIGHT EMITTED AT VARIOUS STAGES
E
D
D (NO LIGHT)
F (VISIBLE)

Fig. 13. Mechanism of positive-polarity creepage

A—Conditions prior to ionization
B—During stepped streamer propagation
C—Completion of first stepped streamer
D—Charge redistribution causes space charge centers
E—Completion of second stepped streamer

are shown. As these space charges build up, electrons will be attracted to them, causing further ionization around them. If a charge center builds up near the channel tip, an extension of the streamer will result. If another charge center occurs in the mid-channel, a branching will result. This is schematically shown at *E* in Fig. 13.

It is now apparent that as long as the gradient near the tips can build up above the critical value for further ionization, the streamer will progress for another step. The whole process then repeats itself. Otherwise it will come to a stop.

In the foregoing explanations of positive and negative creepage, many factors were not considered. Among these were magnitudes of photoionization, effects of the solid insulation, interactions between streamers, and losses of electrons attributable to diffusion and attachments in the liquid and solid structures. Knowledge of these and other factors must come from future work. They are absolutely necessary when one deals with the creepage phenomenon on a quantitative basis.

Conclusions

1. As negative polarity impulse creepage moves over an insulating surface, it propagates in a manner analogous to the propagation of stepped streamers in air similar to natural lightning.

2. Probe measurements indicate that this streamer may carry full-line potential for a considerable distance out over the surface and that this potential rises and falls in a manner explained by the alternate progression of a weak pilot streamer and a brighter stepped streamer through the oil.

3. A positive polarity creepage streamer is shown to propagate by a series of discrete discharges that differ from negative polarity in both modes of propagation and channel potentials.

4. Under the test condition described, the average velocity of propagation for both positive and negative streamers is about 0.1 inch per microsecond.

5. The mechanisms of both positive and negative impulse creepage in oil are postulated in considerable detail.

Appendix. Calculation of Pilot Streamer Characteristics

In Fig. 14, an idealized pilot streamer is sketched. If one assumes that the region surrounding the channel is homogeneous and isotropic, and that the pilot streamer is largely a "minimal" streamer (in that it represents the minimum possible condition for propagation, so that charge density, channel size, and velocity of propagation are minimal), then certain approximate calculations of channel characteristics may be made.

The approximate relationships are

$$q_c = \frac{2\epsilon E_r}{r} \tag{1}$$

$$i_s = q_c K E_s \tag{2}$$

$$R = \frac{1}{2\pi r \epsilon E_r K} \tag{3}$$

$$C = \frac{24}{\log_{10} \dfrac{2h}{r}} \tag{4}$$

$$v_s = K E_{sT} \tag{5}$$

$$V = \frac{\pi r \epsilon E_r \log_{10} \dfrac{2h}{r}\, 10^{12}}{12} \tag{6}$$

where

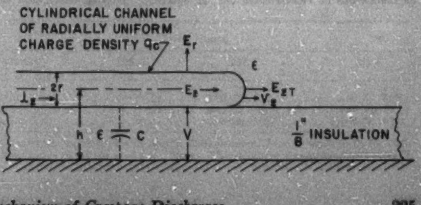

CYLINDRICAL CHANNEL OF RADIALLY UNIFORM CHARGE DENSITY q_c
E_r
ϵ
$2r$
I_s
E_s
E_{sT}
V_s
h
ϵ
C
V
$\frac{1}{8}$" INSULATION

Fig. 14. Idealized pilot streamer characteristics and equations

ϵ = dielectric constant of both oil and solid insulation
E_z = axial gradient within the streamer channel
E_{zT} = axial gradient at the streamer tip
v_s = velocity of streamer tip
q_c = channel charge density
i_z = axial current density

Then if the axial gradient along the channel does not vary drastically, and the radial gradient at the channel walls is the critical breakdown gradient of oil, then equation 1, for channel charge density in coulombs per cubic meter can be derived from Gauss' law. For gradients and radii encountered in oil, a charge density of about 10 coulombs per cubic meter (or roughly 6×10^{12} electrons per cubic centimeter) is obtained. This is about 3×10^{3} times the number of electrons postulated by Schonland[8] in the lightning pilot streamer. It is to be expected that in liquids the election density would be much higher.

Equation 2 is the normal conduction equation from which one can derive total pilot streamer channel currents of a few milliamperes, assuming an electron mobility in oil of 10^{-4} meter per second per volt per meter and axial gradients within the channel somewhat less than the critical oil breakdown gradients at the channel tip.

Equation 3 is derived by combining equations 1 and 2, and for small radii it may reach several thousand megohms per meter or tens of megohms per centimeter. Combining this value with the approximate channel capacitance to ground given by equation 4 as possibly 0.5 micromicrofarad per centimeter, one arrives at a channel charging "time constant" of 10 to 50 microseconds per centimeter of length. This approaches the average propagation time of 4 microseconds per centimeter for the breakdown process.

Equation 5 is a simple expression for velocity of electrons at the channel tip, and gives the approximate value of creepage velocity as discussed in the text if the tip gradient E_{zT} is taken as the breakdown gradient of oil.

Equation 6 is derived from the simple relation $V=Q/C$ where Q is derived from equation 1 and C is the approximate expression of equation 4. This expression also predicts the same order of magnitude of potential to ground as that measured by the probes.

References

1. Dielectric Breakdown of Solids (book), **S. Whitehead.** Oxford University Press, London, England, 1951.

2. Electrical Breakdown of Gases (book), **J. M. Meek, J. D. Craggs.** Oxford University Press, London, England, 1953.

3. Electrical Discharges Over the Surface of Insulators in Transformer Oil, **H. Staack.** *Archiv fur Elektrotechnik*, Berlin, Germany, vol. 25, 1931, pp. 607–30.

4. Impulse Ionization and Breakdown in Liquid Dielectrics, **G. M. L. Sommerman, C. J. Bute, E. L. C. Larson.** *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 147–55.

5. Surface Discharges on Oil-Immersed Insulation, **A. Imhof.** *Paper No. 220*, CIGRE, Paris, France, 1937.

6. Propagation Mechanism of Impulse Corona and Breakdown in Oil, **T. W. Liao, J. G. Anderson.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 641–48.

7. Lightning to the Empire State Building, **K. B. McEachron.** *AIEE Transactions*, vol. 60, Sept. 1941, pp. 885–90.

8. The Pilot Streamer in Lightning and the Long Spark, **B. F. J. Schonland.** *Proceedings of the Royal Society*, London, England, vol. 220, Series A, 1953, pp. 25–38.

9. The Development of the Spark Discharge, **T. E. Allibone, J. M. Meek.** *Ibid.*, vol. 166, Series A, 1938, pp. 97–126.

10. Mechanism of Dielectric Breakdown of Liquids, **V. S. Komelkov.** *Comptes Rendus (Doklady)*, Academie des Sciences de l'URSS, Moscow, USSR, vol. 48, no. 4, 1945, pp. 268–70.

11. Electrical Breakdown of Solid and Liquid Insulations, **A. von Hippel.** *Journal of Applied Physics*, New York, N. Y., vol. 8, 1937, p. 815.

◆

Discussion

G. M. L. Sommerman (Westinghouse Electric Corporation, East Pittsburgh, Pa.): By the use of probe electrodes, the authors have added another valuable experimental technique to those previously reported (reference 6 of the paper) for investigating the propagation of impulse streamers over a solid-fluid interface. This permits data to be obtained on streamer propagation velocities (complementing those obtainable from the Boys camera and other photographic means) but mainly it provides data for obtaining information on streamer gradients as a function of time and distance along the streamer. From the last, it is possible to estimate streamer charge distribution and current density as functions of time and distance, as the authors have shown.

In the calculations of streamer current densities as described in the Appendix and Fig. 14, it is evident that assumptions have to be made for the gradients normal to the streamer and for electron mobility in the immersant. The latter is especially difficult to estimate reliably. It is suggested that estimated values of streamer currents and densities could also be obtained from current-time oscillograms for the whole streamer array and an analysis of the photographic record of the array. Such oscillograms were included in the previous paper (see reference 6), but quantitative co-ordinate scales were not reported. It would be helpful if the authors would report in their closing discussion some of these quantitative total streamer-array current values as a function of time.

With respect to the use of long-tailed impulse-voltage waves by the authors, the longer streamer lengths obtained with such waves compared with those for $1\frac{1}{2}\times40$-microsecond waves confirm results obtained by Bute, Larson, and the writer with waves of about ten times the standard tail (reference 4 of the paper). For waves of this type, reference 4 also gives results with waves chopped at various times to permit taking photographic records of streamer formation in liquids as a function of time of application of voltage. Such data, of course, overcome the time-dependence limitation of the simple Lichtenberg-figure or photo-paper technique mentioned by the authors in their second paragraph. The chopped-wave photo-patterns do yield quantitative information of streamer propagation velocities, although they do not give all of the information obtainable from the Boys camera used by the author. The photographic records whether integrated or time-dependent are still our best means for determining streamer densities as a function of spatial co-ordinates in an insulated structure, although, once known, the streamer paths may be evaluated further by the probe technique described in the paper.

J. G. Anderson and **T. W. Liao**: We appreciate Dr. Sommerman's discussion and are happy to see the check between our results and those obtained by Mr. Bute, Mr. Larson and Mr. Sommerman on the wave dependence of streamer lengths.

Dr. Sommerman's suggestion on estimating stepped-streamer currents was tried in our previous investigation (reference 6 of the paper) in which the streamer current oscillograms were presented. However, because of the uncertainty of the amplifier gain and the oscilloscope response for the exceedingly fast streamer pulses, the values of currents were not reported in that paper. A rough estimate, based on a calibration of the measuring circuit with simulated pulses, gave a current of approximately one ampere for a streamer 3 inches long. Somewhat smaller current values were obtained for the shorter stepped-streamers. With these estimated values a current density of about 500 amperes per square inch would be obtained for a stepped-streamer channel of 50-mil diameter. Similar values of stepped-streamer current densities can be obtained by estimating the currents required to charge the channel capacitance to full voltage in the streamer formation time of less than 1 microsecond.

It is difficult to use pulse current amplitudes to compute the pilot-streamer channel characteristics because the pilot appears to propagate continuously. Thus the computations and assumptions of Fig. 14 and the Appendix become necessary.

A Method for Evaluation of Thermal Stability of Magnet Wire Enamel

C. G. CURRIN
ASSOCIATE MEMBER AIEE

J. F. DEXTER
MEMBER AIEE

Synopsis: A method for evaluating the thermal stability of magnet wire enamel has been developed using the dielectric strength of the enamel film as a criterion. This test method was designed for accuracy, simplicity, and economy in time, labor, materials, and equipment. Standard National Electrical Manufacturers Association (NEMA) dielectric strength twists[1] are dipped in an electrical impregnating varnish and used as test samples. To achieve sufficient accuracy, a large number of these twists are included in an enamel evaluation program. The twists are placed in aging ovens at various temperatures. Periodically the samples are removed and tested by applying a high voltage. The average aging time required to reduce the dielectric strength of the twists to a preselected value is considered the life of the enamel. The relative life of the enameled wire insulation system over a wide temperature range may then be predicted from the life of the samples determined at several temperatures. Studies have been made of this test method under different sample preparation methods, failure criteria, and aging conditions. Results determined by this test are compared with motor-aging test results on the same enamel-impregnating varnish insulation systems. This comparison shows that this is a convenient functional test method that can be correlated with the results of a more complicated motor test program.

TESTS are developed to predict the behavior of a material in service. The purpose of the twist-aging test is to predict accurately the life of a magnet wire enamel-impregnating varnish insulation system in service with a minimum expenditure of time, labor, material, and equipment. This is accomplished by using a number of easily constructed samples subjected to accelerated aging and functionally evaluated.[2]

The test is not designed to replace motor or similar functional equipment tests. Instead, it is a laboratory test method designed to screen a variety of insulation systems. A motor or similar aging test program is then used for confirmation of the most desirable system. This permits the use of an inexpensive test method for screening and the more expensive and more accurate tests for final evaluation.

In evaluating this screening test method, the following requirements were considered:

1. Results must have a high correlation with service.
2. Results must have high statistical accuracy.
3. Results must be reproducible.
4. Failure criteria must be functional.
5. Time required for failure must be as short as possible.
6. Application of conditioning must be simple.
7. Number of conditioning factors must be kept to a minimum.
8. Evaluation of samples must require a minimum of time, material, skill, and equipment.
9. Preparation of samples must require a minimum of time, material, skill, and equipment.
10. Samples must be consistently fabricated, replacing skill with tools wherever possible.
11. Test must be responsive only to factors determining the life of an insulation system in service and in the same order of magnitude.

Factors Considered

Two major factors affecting the life of insulating systems are thermal degradation and dielectric fatigue. To simplify the test procedure, dielectric aging has not been included in this test. Other factors such as moisture, chemical contamination, and mechanical vibration usually do not in themselves cause failure of an insulation system, but are factors that result in failure only after the system has been weakened by thermal deterioration.

The test sample is a standard NEMA dielectric strength twist of the enameled wire, reinforced for mechanical strength, identified, and impregnated with an electrical insulating varnish.[1] The samples are aged at several temperatures. The life-temperature relationship is then established over a temperature range. This relationship[3] is the goal of the test and from these data the life at lower temperatures can be predicted.

Since the primary purpose of magnet wire enamel is to serve as a dielectric barrier, the over-all effectiveness of the enamel may be determined from the dielectric properties. Any fault in the dielectric barrier of the enamel film which would cause an apparatus to fail may be detected by a decrease in the dielectric strength of the enamel. Dielectric strength is one dielectric property which is easily measured. This is the functional and practical basis for selecting dielectric strength as a criterion for enamel failure.

For simplicity, the test voltage is applied for only 1 second and is applied periodically during the aging of the twists. When all the samples have failed the test voltage, the average life at the particular aging temperature is determined. The life-temperature characteristic may then be plotted.

Some of the factors of this test method which have been studied are:

1. Sample preparation methods.
2. Effect of varnish impregnation.
3. Effect of thickness of varnish film.
4. Failure criteria.
5. Effect of moisture conditioning.
6. Effect of wire size.
7. Reproducibility of results.
8. Effect of thermal shock.
9. Aging temperatures.
10. Number of samples required.

Studies of these factors were made by

Table I. Modified Silicone Enamel Life Versus Impregnation

Enamel Coated to Heavy Build on No. 18 Copper Wire. Twist Samples Impregnated with One Coat of Varnish. Samples Tested with a Failure Criterion of 2.3 Kv (30% of Original Dielectric Strength)

Impregnating Varnish	Life,* Hours			
	300 C	275 C	250 C	225 C
Organic	9.6 (11, 15%)	40 (6, 10%)	209 (10, 20%)	485 (21, 31%)
Silicone A	12.8 (5, 10%)	66 (9, 17%)	246 (3, 6%)	1,387 (10, 20%)
Silicone B	12.8 (10, 18%)	65 (5, 9%)	222 (9, 17%)	1,160 (9, 16%)

*Life given as A, B, and C% where A = average life in hours; B = tolerance of the average in % for 99% probability; C = standard deviation in % of the average. Note: This footnote also applied to Tables II through V.

Paper 55-46, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted June 30, 1954; made available for printing November 26, 1954.

C. G. CURRIN and J. F. DEXTER are with the Dow Corning Corporation, Midland, Mich.

Table II. Synthetic Organic Enamel Life Versus Failure Criterion

Enamel Coated to Heavy Build on No. 18 Copper Wire. Twist Samples Impregnated with One Coat of a Silicone Enamel

Failure Criterion	Life, Hours 300 C	275 C	250 C	225 C	200 C
3.1 kv, 50% of original dielectric strength	9.9 (27, 51%)	22.5 (21, 40%)	40.0 (32, 63%)		507 (17, 34%)
1.9 kv, 30% of original dielectric strength	13.4 (16, 29%)	35.9 (10, 19%)	76.4 (12, 23%)	204 (16, 30%)	
2.0 kv	15.4 (11, 16%)	47.9 (12, 18%)	85.8 (11, 16%)	186 (0, 0%)	
1.0 kv	19.7 (8, 12%)	48.5 (16, 24%)	109.0 (10, 15%)	392 (12, 17%)	
0.5 kv	18.6 (11, 17%)	44.7 (12, 18%)	130.5 (11, 15%)	540 (13, 20%)	

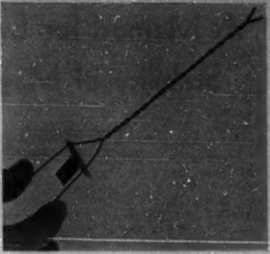


Fig. 1. Typical test sample

Table III. Comparison of Results on Polyvinyl Formal Enamel by Twist-Aging Test and Motor-Aging Tests

Test	Temperature, C	Life, Hours	Remarks
Twist-aging test	300	1.5	Failure criterion of 30% of original dielectric strength
	275	7.8	
	250	12.5	
	225	30.6	
	200	197	
Reversing-motor test	225	80	System B, reference 8
	200	250	
	175	700	
Motor test	200	154	System N, reference 6
	160	1,869	
	135	8,290	
Coil test	260	13	Reference 9
	230	30	
	200	110	
	180	350	
Motor test	200	140	Reference 7
	160	980	
	135	3,240	

Table IV. Comparison of Results on Modified Silicone Enamel by Twist-Aging Test and Motor-Aging Test

Enamel Coated to Heavy Build on No. 18 Copper Wire. A Silicone Varnish Used for Impregnation

Test Method	Life, Hours 300 C	275 C	250 C	225 C	200 C
Motor-aging test*		58, 60, 68	235, 411	841, 618, 1,828, 1,967	2,354, 3,160, 5,4
Twist-aging test†	13.4 (6, 10%)	68 (4, 6%)	245 (4, 6%)	1,694 (0, 0%)	

*Reversing motor test, 284 frames, 1,200 rpm, 3 phases, 440 volts. See reference 10 for details.

†Twists evaluated with failure criterion of 1.0 kv.

Table V. Elevated Temperature Life of Various Wire Enamels

All Enamels Coated to Heavy Build on No. 18 Copper Wire. Samples Impregnated with One Coat of a Silicone Varnish; Failure Criterion, 2.0 Kv

Wire Enamel	Life, Hours 300 C	275 C	250 C	225 C
Oleoresinous enamel	6.7 (32, 40%)	19.5 (15, 18%)	47.5 (13, 23%)	91.2 (14, 18%)
Polyvinyl formal enamel	2.5 (41, 61%)	8.6 (24, 35%)	13.0 (16, 25%)	28.6 (19, 29%)
Synthetic organic enamel	13.4 (16, 29%)	35.9 (10, 19%)	76.4 (12, 23%)	204.0 (16, 30%)
Modified silicone enamel	11.5 (4, 6%)	63.0 (6, 9%)	260.0 (5, 7%)	1,384.0 (11, 17%)

determining the life-temperature curve of several magnet wire enamels, varying one factor for each test. Polyvinyl formal, a modified silicone enamel, and a synthetic organic enamel were evaluated. Approximately 20 samples were aged at each temperature ranging from 200 to 300 degrees centigrade (C). The samples were periodically removed from the aging ovens and evaluated by being subjected to the test voltage for 1 second. The average life of each group of samples at each temperature was determined as well as the standard deviation and the statistical accuracy[4] of the data. These data are given in Tables I through V. The data were also plotted to determine the life-temperature characteristics for each study. These curves were determined by plotting the average values and constructing the curves for a maximum probability of accuracy. To simplify interpretation of the curves shown, the individual points and their probability limits as given in the tables are not shown.

Sample Preparation Method

The standard NEMA dielectric twist was selected because of its simplicity, low cost, and functionality. Fig. 1 shows the completed test sample. It was found necessary to limit the bending of the leads where they enter the twisted section of the sample. A jig was designed to assure uniform bending of the leads. The spacer between the leads reduces the bending moment on the leads where they enter the twisted section. The spacer reduces mechanical damage to the insulation at this critical spot resulting from the handling of the sample during the aging periods. The spacer also serves as a support for an identifying label on each twist. Spacers must be of a material equal to or better than the wire enamel in thermal stability.

Each twist is impregnated with a suitable electrical varnish. The impregnating varnish is cured at a temperature which has a negligible effect on the wire enamel.

After impregnation, all samples of one test series are grouped together in aluminum radio chassis which are convenient oven trays. These trays are arranged so that the oven air may flow freely over all the samples. Three to 5 minutes are allowed for the samples to reach the aging temperature in forced air ovens.

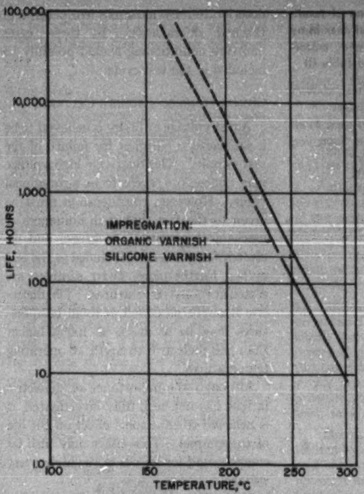


Fig. 2. Effect of impregnation on life of modified silicone enamel. Enamel was coated to heavy build on no. 18 copper wire. Samples were impregnated with one coat of varnish and tested with a failure criterion of 2.3 kv (30% of original dielectric strength). See Table I

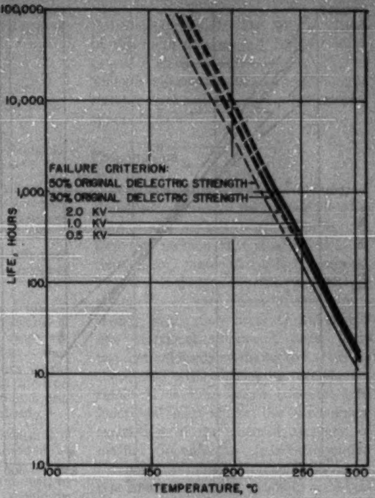


Fig. 3 (right). Effect of failure criterion on life of modified silicone enamel. Enamel was coated to heavy build on no. 18 copper wire and impregnated with one coat of silicone varnish. See Table II

EFFECT OF IMPREGNATION

Evidence has been found to show that the impregnating varnish has a direct effect on the life of a wire enamel. This is perhaps because of the varnish acting as an oxygen barrier, thus affecting the oxidation rate of the wire enamel. The varnish also provides mechanical protection. Because this test method is designed to be functional, the samples are impregnated with a varnish which would normally be used with the enamel. Although this method is primarily a test of the wire enamel, it is more realistic when it includes the effects of the immediate environment, which is the varnish. Fig. 2 shows the effect of two varnishes on one common wire enamel. The curve shows that the life of the enamel is materially affected by the impregnation.

EFFECT OF THICKNESS OF VARNISH FILM

An investigation was made to determine whether the inpregnating varnish thickness has an effect on the enamel life. The life-temperature curve of samples with one dip in varnish was compared with the life of similar samples that were given two dips. There is little difference between the life of the two samples. Although the varnish thickness on the surface of the samples varies from 0.003 inch with one coat to 0.006 inch with two coats, the varnish thickness between the two twisted wires of the sample remains constant.

FAILURE CRITERIA

Another important consideration in any test method is the failure criterion. It should be a functional property that can be accurately and economically measured. Studies were made of failure criteria of two types: a percentage change in dielectric strength, and a predetermined minimum dielectric strength. These studies show that the failure criterion has a significant effect on the test results. Fig. 3 shows the effect of the failure criterion on tests of one wire enamel. With a failure criterion of 50 per cent (%) of the original dielectric strength after impregnation, the test is completed in the shortest time, but the results are not realistic because enamels with a higher initial dielectric strength are tested under more severe conditions.

The testing will be simpler and the criterion will be more functional if the failure criterion is a fixed dielectric strength that takes into account the stresses to which the dielectric barrier is subjected under service conditions. If the dielectric strength selected as a failure criterion is too high, the enamel will be subjected to abnormally severe stresses and the test results will be less accurate. The effects of the failure criterion on the accuracy of the test results are indicated in Table VI. It is apparent that the statistical accuracy of the test results increases as the voltage criterion is decreased.

MOISTURE CONDITIONING

Moisture is an environmental factor that should be considered in any functional test of electrical insulating materials. The effects of moisture conditioning of the twist sample has been investigated. One group of samples was subjected to a humidity cycle during the aging period. Modified silicone-enameled wire samples impregnated with a silicone

Table VI. Average Standard Deviation, % of Mean

Criterion	300 C	275 C	250 C	225 C	200 C	All
50% of original dielectric strength	50	44	55		28	47
30% of original dielectric strength	25	33	27	27	21	27
2.0 kv	27	27	21	14		22
1.0 kv	19	27	24	16		21
0.5 kv	18	19	21	17		18

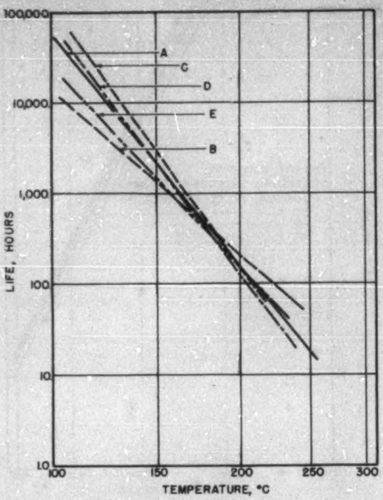


Fig. 4. Comparison of results on polyvinyl formal enamel by twist-aging test and motor-aging tests. See Table III

A—Twist-aging test, class-A insulating system
B—System B, from Leape et al[4]
C—System N, from Scott and Thompson[5]
D—From Brancato[6]
E—From Hermann[7]

varnish were aged at 250 C. The 25 samples were cycled 24 hours at the aging temperature and then 24 hours at 100% relative humidity at 25 C. The samples were tested at 2.0 kv after each aging cycle and the voltage was reapplied 10 minutes after each humidity cycle.

All but 3 of the 25 samples failed after the thermal aging cycle, rather than after the humidity cycle. This indicates that for the insulation tested the humidity cycle was not an important factor in determinging failure points when a relatively high-voltate failure criterion was used. Furthermore, the humidity cycle greatly complicated the test procedure and increased the time required to complete the test. Some materials may be deteriorated by moisture irrespective of thermal degradation. In these cases moisture conditioning should possibly be included in the test cycle.

Conclusions on Failure Criteria

A test voltage of 1.0 kv is believed to be a satisfactory criterion for failure of the test sample. This relatively high voltage may appear too severe to be a functional value. However, consideration must be given to the fact that such influences as mechanical vibration, moisture, and dielectric fatigue are not included in the test cycle. Furthermore, twist samples are tested at room temperatures. The dielectric strength measured at room temperatures may be as much as 100% higher than the dielectric strength at operating temperatures.

Although the importance of dielectric fatigue has not been fully investigated, it is believed to have some effect on the life of the enamel. This effect may also be compensated for by use of the higher test voltage.

The twist-aging test must be sensitive to any physical discontinuities of the enamel film. This requires a voltage stress between the twisted wires greater than the dielectric strength of air with the same electrode spacing. A criterion voltage of 1.0 kv appears to be necessary to overcome the dielectric strength of the air where moisture is not introduced into the test procedure, and to compensate for the elimination of electrical fatigue as a deteriorating influence. Furthermore, data obtained using this value as

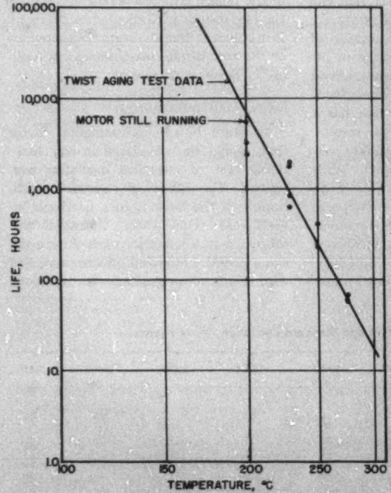
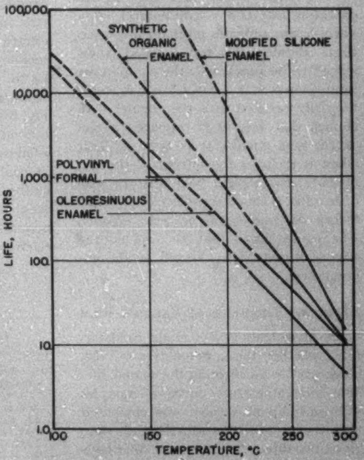


Fig. 5 (left). Comparison of results on modified silicone enamel by twist-aging test and a motor-aging test. Points indicate lives of 10-horsepower 1,200-rpm 3-phase 440-volt 284-frame induction motors. See Table IV

Fig. 6 (right). High-temperature life of various wire enamels. Enamels coated to heavy build on no. 18 copper wire. Samples impregnated with one coat of silicone varnish. Samples tested with a failure criterion of 2.0 kv. See Table V



the failure criterion correlates well with published motor test results as described later.

Wire Size

For convenience in fabricating and handling, the samples should be made of magnet wire of approximately no. 18 American Wire Gauge. Tests of a modified silicone enamel on no. 20 and no. 18 copper wire show no significant difference in resultant enamel life.

Reproducibility

One sample of modified silicone enameled wire was evaluated twice by the twist-aging test, starting at different times. There is a high probability that the results of the tests were identical, considering the statistical accuracy of the data. This is evidence that the results obtained with the twist-aging test can be reproduced.

Thermal Shock

Twist-aging test samples are subjected to thermal shock each time the samples are placed in the ovens and then removed for testing. Tests were conducted to determine the effect, if any, of thermal shock.

A large number of samples of modified silicone-enameled wire were aged at 250 and 300 C. Half the samples at each temperature were subjected to thermal shock similar to other twist-aging tests. The other half were not subjected to thermal shock except for the initial entrance into the oven and removal from the oven for destructive dielectric testing. The dielectric strength of the samples subjected to thermal shock did not vary appreciably from that of the samples not subjected to thermal shock. This was observed during the life of the samples at both aging temperatures. Thus, from this study, thermal shock does not appear greatly to affect the results.

However, an observation made when testing many lots of samples shows some effect of thermal shock. Samples which were not tested during the initial 50% of their life, and were therefore less frequently exposed to thermal shock, had about 10% longer life than samples tested by the standard procedure, which is to test the samples at intervals of 5 to 10% of the expected life throughout the life of the sample. This procedure reduces book work and simplifies the testing of a large number of enamels. Thus, if a prescribed standard procedure is followed, the effect of thermal shock will be constant on all samples; and if the testing intervals deviate from the standard by as much as 50%, then the results may vary approximately 10%.

Aging Temperature

Several principles are involved in selecting aging temperatures:

1. Aging at high temperatures decreases the time required for test completion.

2. Aging at high temperatures where the life is extremely short decreases the accuracy of the test. Results obtained on samples that fail in 20 hours or less are not reliable.

3. Accuracy of the slope of the life-temperature curve is improved as the range of test temperatures is increased.

4. Life expectancy at operating temperatures can be extrapolated most accurately when the lowest aging temperature approaches the operating temperature. Furthermore, lower aging temperatures are desirable because the slope of the life-temperature curve may change with temperature, making extrapolation unreliable.

5. The number of aging temperatures is limited by the practical consideration of the cost of separate oven and control equipment for each aging temperature. However, at least three aging temperatures should be used to determine the life-temperature characteristic of the insulating material.

Number of Samples Necessary

The statistical accuracy of the test is dependent upon the following:

1. The number of samples and the consistency of the failures of each group of samples at each temperature. Ten samples will usually give results of adequate statistical accuracy.

2. Accuracy of temperature control. A tolerance of not more than ±1C is recommended.

3. Range in the aging temperatures.

Correlation of Twist-Aging Data to Motor Test Data

None of the manual operations of the test procedure[5] require any appreciable skill, and thus the test may be performed by unskilled personnel. The twist aging method is of value only if there is close correlation with the results obtained by the more functional motor tests. Motor-aging tests have been recognized as an accurate basis for estimating insulation life.

Fig. 4 shows the correlation between the twist-aging test of polyvinyl formal insulated magnet wire and published motor-aging tests,[6-8] and one coil-aging test.[9] This figure shows that the twist-aging test correlates as closely with the various motor tests as these motor tests correlate with one another. Fig. 5 shows close correlation between the twist-aging test and a motor-aging test[10] on a modified silicone enamel. The data obtained by this method on several different magnet wire enamels are shown in Fig. 6. This curve shows the difference in thermal stability that may be expected of various types of enamel.

Conclusions

The twist-aging test has been designed to predict accurately the service life of a magnet wire enamel by a simple laboratory test method that is economical in time, labor, materials, and equipment. This test is designed for the preliminary evaluation of wire enamels before proceeding to the more accurate and expensive functional equipment tests. This test method has been shown to have close correlation with various motor test results. It is therefore a practical and functional method for the preliminary evaluation of the thermal stability of magnet wire enamels. It is not intended to replace the more comprehensive evaluation of magnet wire enamels in combination with other insulating materials in electric machines.

References

1. NEMA Standards for Magnet Wire. *Publication No. 46-120*, National Electrical Manufacturers Association, New York, N. Y., *MW 24-50*, 1946, paragraph A.

2. Functional Evaluation of Insulation Materials, K. N. Mathes. *AIEE Transactions*, vol. 67, pt. II, 1948, pp. 1236–40.

3. Electrical Insulation Deterioration Treated as a Chemical Rate Phenomenon, Thomas W. Dakin. *Ibid.*, pt. I, pp. 113–22.

4. Statistical Methods (book), G. W. Snedecor. Iowa State College Press, Ames, Iowa, fourth edition, 1946, chap. 3.

5. Thermal Evaluation of Insulating Materials. *AIEE Special Publication S-61*, "A Method for Evaluation of Thermal Stability for Magnet-Wire Enamel," C. G. Currin, J. F. Dexter. June 1954, pp. 19–26.

6. Temperature-Aging Tests on Class-A-Insulated Fractional-Horsepower Motor Stators, J. A. Scott, B. H. Thompson. *AIEE Transactions*, vol. 61, July 1942, pp. 499–501.

7. Motor Insulation as Measured by Accelerated Test and Dielectric Fatigue, C. J. Herman. *Ibid.*, vol. 72, pt. III, Oct. 1953, pp. 986–91.

8. A Method of Evaluating Insulation Systems in Motors, C. B. Leape, J. McDonald, G. P. Gibson. *Ibid.*, Aug., pp. 793–99.

9. Nondestructive Testing of Insulation-Emanuel L. Brancato. *Ibid.*, pt. I, May, pp. 203–07.

10. Thermal Endurance of Silicone Magnet Wire Evaluated by Test, W. J. Bush, J. F. Dexter. *Ibid.*, vol. 73, pt. III, Aug. 1954, pp. 1005–10.

◆

Discussion

G. M. L. Sommerman (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The repeatability on wire enamel of these life test results, as well as those obtained by other organizations, is very good indeed. This is noteworthy because insulation life results usually show greater deviations as the insulation thickness is decreased to the film forms characteristic of magnet-wire enamel insulation. The accuracy attained is due to the careful planning given to sample preparation, life-test procedures and test criteria which have been contributed by the authors and other workers engaged in such investigations.

In regard to the thermal-aging tests including the high-humidity cycle, it is rather surprising that as many as 88 per cent of the samples tested did not fail after the humidity cycles. Perhaps the humidity, while very close to 100 per cent, was not quite sufficient to cause condensation, which would have permitted free moisture to be drawn by surface-tension forces into any crevices developed in the enamel by thermal aging. It has been suggested that a short water-immersion test might be applied to the samples in lieu of a high-humidity test. This would be relatively simple, less time-consuming, and yet might be more positive than the 24-hour "100-per-cent" humidity test with its inevitable tolerances. In fact, the immersion tests could be run immediately after the dry voltage-withstand tests; also they might be applied to only half the samples tested until experience is obtained with this added feature.

A few preliminary experiments on water-immersed aged enameled wire samples by F. A. Sattler and the writer indicate that this procedure can detect cases where fissures in the enamel walls of the two twisted wires are separated by several times the double enamel thickness. This may be important where equipment is operated at light loads in a cold atmosphere and then brought to a warmer high-humidity atmosphere, especially if voltage transients may be applied to coil terminals.

C. G. Currin and J. F. Dexter: The authors wish to thank Dr. Sommerman for his comments. Following the first presentation of this paper at the 1954 Winter General Meeting, a task force was set up under the sponsorship of the Basic Sciences Subcommittee on Dielectrics to develop a test procedure for the evaluation of the thermal stability of enameled wire. The method described in this paper was selected by the task force and a co-operative test program was started to obtain experience using this method. Six laboratories participated in the program, in which two different types of enameled wire were evaluated. The results obtained by each laboratory agreed very well. Subsequent work by the task force has resulted in a test procedure that has been presented to the AIEE for approval for trial use.

In regard to the question of high relative humidity conditioning versus water immersion, the author's recent work also shows that a low-voltage dielectric test after water immersion will detect cracks and flaws in the enameled films more readily than a high relative humidity conditioning. Dr. Sommerman's suggestions should be considered in any revision of the proposed test procedure for the evaluation of the thermal stability of enameled magnet wire.

Insulation Aging Characteristics as Measured by Shearing Modulus

L. C. WHITMAN
MEMBER AIEE

A. L. SCHEIDELER
ASSOCIATE MEMBER AIEE

Synopsis: A method is described for determining the effect of thermal aging on the shearing modulus of flexible sheet insulations by nondestructive tests using a torsion pendulum technique. The test sample, in tape form, is simple and easily prepared. Shearing modulus of the material is determined with the tape functioning as the stressed element in a simple torsion pendulum. Equipment for sample preparation and testing is inexpensive and compact, allowing investigations of many materials with minimum expense and effort. The data obtained by this method of test are of particular value and interest in determining the effect of time and temperature on an important mechanical property of flexible sheet materials. Since tests may be made as a function of aging on a given sample, the technique is very useful in establishing deterioration rates of the shearing modulus for a given insulation. Data have been obtained upon materials of widely different temperature classifications in order to demonstrate the usefulness of the technique.

NUMEROUS methods have been proposed for determining the thermal aging characteristics of electrical insulations. The rate of change of some mechanical or electrical property of the material as a function of time and temperature is generally utilized in specifying life criteria. The measurements used in evaluating these properties may be placed in two broad categories: destructive tests and nondestructive tests.

Such destructive measurements as dielectric strength, tensile strength, and bursting strength have been of demonstrated value in evaluating the aging characteristics of many materials. Generally, however, large numbers of samples are required and it is difficult to determine deterioration rates accurately since only one test may be made on a single sample.

Such nondestructive electrical measurements as insulation resistance and power factor have been used with some success in thermal aging studies. Owing to the difficulties of measurement little attention has been given to similar tests of a mechanical nature. It may be shown, however, that a simple, nondestructive mechanical test using shearing modulus as a criterion is very useful in the study of flexible sheet insulations. This particular property has not been employed previously since no simple method of measurement had been developed. It is believed, however, that it is of major importance since such a basic characteristic of apparatus as the short-circuit strength is closely allied with insulation shear.[1]

Measurement of Shearing Modulus With the Torsion Pendulum

The shearing modulus (or coefficient of rigidity) of a vertical round wire may be determined quite simply by measuring the period of oscillation of an inertia disk attached to its lower end. The mathematics of such a system are relatively simple and the shearing modulus of any isotropic material that can be drawn or formed into a round rod, wire, or thread may be determined easily. Most insulations, however, are flat, thin sheets or tapes, or laminated sections thereof. If a torsion pendulum is constructed of such material, the restoring force producing oscillation of the inertia disk or bar results from the torque necessary to twist a thin rectangular bar.

This torque may no longer be simply expressed and has been the subject of many investigations. Solutions were proposed as early as 1855 by DeSaint-Venant,[2,3] by Buckley[4] in 1914, and by others, but no complete solution appears to have been published until Timoshenko[5] in 1941. This is abstracted in Appendix I to show the effects considered and is later utilized in relating shearing modulus to torsion pendulum period.

Paper 55-35, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 15, 1954; made available for printing November 16, 1954.

L. C. WHITMAN and A. L. SCHEIDELER are with the General Electric Company, Pittsfield, Mass.

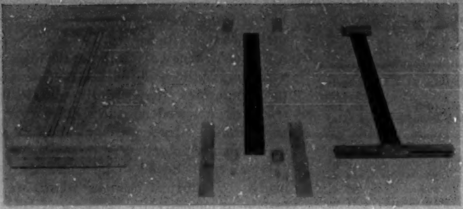


Fig. 1 (left). Torsion pendulum test sample and alignment fixture

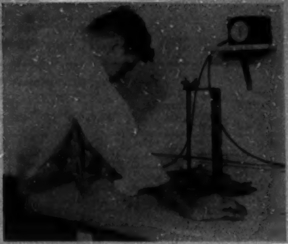


Fig. 4. Test equipment for determining torsion pendulum period

Test Equipment

The equipment for the preparation of samples and testing using the torsion pendulum technique is simple and inexpensive. Except for the aging ovens all the necessary components are illustrated in Figs. 1 through 4. The test piece shown in Fig. 1 consists simply of a suitable length of material (12 inches has proved satisfactory) with metal inertia bars clamped at each end. The inertia bars consist of pairs of steel bars 1 inch by 1/8 inch, the heavier being 6 inches long and the lighter 2 inches long. Note that both sets of these bars are attached to the strip under test and by inverting the test piece either set of inertia bars can be used, the inactive set serving as the top support. Strips of silicone rubber attached to the bars are used to grip the test strip uniformly and to avoid crushing the material.

Care should be taken in assembling the test piece to maintain perpendicularity between the inertia bars and the strip and to have the central axis of the inertia bar coincide with the central axis of the test strip. This may be done by using the simple alignment fixture shown in Fig. 1. A sample rack such as is shown in Fig. 2 is necessary to hold the samples during aging. It is important to give the samples sufficient support to prevent warping or distortion during aging but not to have them under excessive strain. This is accomplished by supporting the strip in a vertical position with a longitudinal stress supplied by the inertia bar. The inertia bar is positioned in slots with small clearance (1/16 inch) to prevent turning or excessive motion due to air currents when aging. It is also important, during aging in the oven, to keep the sample rack level so that no binding or uneven stresses are produced between the top clamping member and the strip.

A timing fixture is necessary which provides true perpendicularity of support for the test piece and some adjustment for minor twisting of the test piece due to aging. The timing fixture shown in Fig. 3 provides an adjustable turntable which may be aligned with the normal relaxed position of the test sample. This relaxed position, rather than the absolute angle referred to top clamp, is taken as the zero angle of twist. The inertia bar is then displaced the specified angle α from this relaxed position and allowed to oscillate. The period of either single or multiple cycles of oscillation may then be timed.

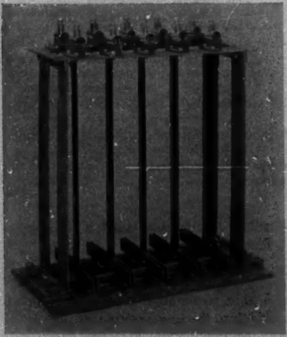


Fig. 2. Rack to hold samples and maintain alignment during aging

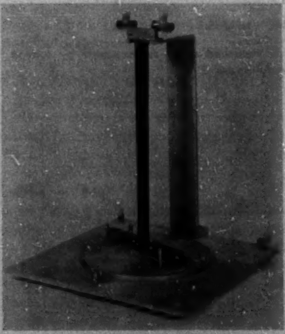


Fig. 3. Timing fixture permitting alignment in relaxed position and indexing of initial angle of twist

Experimental Procedures

Five samples of each material were used at each aging temperature. The period was recorded as the average of five oscillations and five determinations were made for each sample during each measurement. The period was determined before aging, at room temperature and humidity. The samples were then placed in an air recirculating oven and aged at appropriate temperatures. The test pieces were removed from the oven at suitable intervals and measurements of the period were made after cooling for 1 hour at room temperature.

Materials Investigated and Results

The thermal aging characteristics of a number of materials as described in Table I have been investigated using the torsion pendulum technique. The effect of prolonged aging at accelerated temperatures on the period of the pendulums assembled utilizing these materials are illustrated in Table II and Figs. 5, 6, and 7. Tests on black varnished cloth at 175 degrees centigrade (C) and 150 C and on polyethylene terephthalate (My-

Table I. Materials Investigated

Material	Nominal Thickness, Inches	Figure
Black varnished cloth	0.007	5, 8, 9, 11, 12
Silicone resin impregnated glass cloth	0.010	6
Silicone rubber impregnated glass cloth	0.005	6
Polyethylene terephthalate (Mylar)	0.008	7, 8, 10

Table II. Effect of Aging on Shearing Modulus

Black Varnished Cloth: Nominal Thickness 7 Mils, Width 1 Inch

Temperature, C	Hours, Aging	Lighter Pendulum No. 1* T_1, Seconds	Lighter Pendulum No. 1* U_1, 10^{-3} Inch-Pounds	Heavier Pendulum No. 2† T_2, Seconds	Heavier Pendulum No. 2† U_2, 10^{-3} Inch-Pounds	Shearing Modulus G_0, 10^4 Lb/In²
175	0	1.78	0.669	7.62	0.815	12.3
	1	1.56	0.871	6.10	1.27	13.6
	3	1.38	1.113	5.60	1.508	18.7
	6	1.28	1.294	5.28	1.697	22.4
	12	1.14	1.631	4.64	2.197	27.6
	18	1.10	1.752	4.50	2.336	29.9
	33	1.06	1.886	4.34	2.511	32.3
150	0	1.50	0.941	5.44	1.60	12.0
	1	1.66	0.770	6.36	1.17	11.5
	3	1.54	0.895	6.02	1.305	14.1
	5	1.46	0.995	5.88	1.37	16.4
	10	1.36	1.15	5.52	1.55	19.5
	16	1.30	1.255	5.28	1.70	21.2
	21	1.20	1.47	5.06	1.85	26.4
	37	1.16	1.58	4.72	2.12	26.9
	61	1.10	1.75	4.48	2.37	29.4
	87	1.04	1.96	4.06	2.88	30.4
135	0	1.76	0.684	5.96	1.332	6.7
	1	1.86	0.613	6.38	1.162	6.4
	4	1.82	0.640	6.28	1.199	6.9
	6	1.76	0.684	6.16	1.246	7.7
	10	1.70	0.734	6.04	1.297	8.8
	15	1.64	0.788	5.94	1.341	10.0
	33	1.50	0.942	5.66	1.476	13.5
	39	1.48	0.968	5.58	1.519	13.8
	47	1.44	1.022	5.50	1.564	15.1
	56	1.38	1.113	5.28	1.697	16.5
	68	1.34	1.180	5.22	1.736	18.2
	92	1.32	1.217	5.06	1.848	18.1
	142	1.24	1.378	4.90	1.97	22.0
	206	1.22	1.425	4.78	2.07	22.3

* For the lighter pendulum no. 1, $m_1 = 1.828$ ounces. $I_1 = 1.07 \times 10^{-4}$ inch-pound-seconds².

† For the heavier pendulum no. 2, $m_2 = 5.098$ ounces. $I_2 = 23.96 \times 10^{-4}$ inch-pound-seconds².

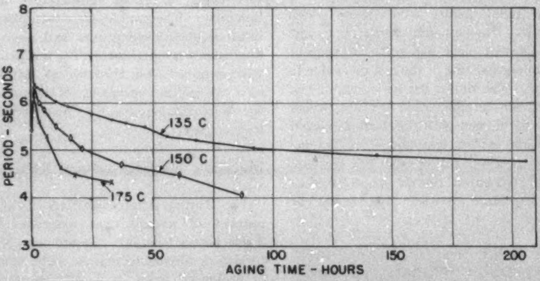


Fig. 5. Mechanical aging characteristics of black varnished cloth as measured by torsion pendulum, using pendulum no. 2

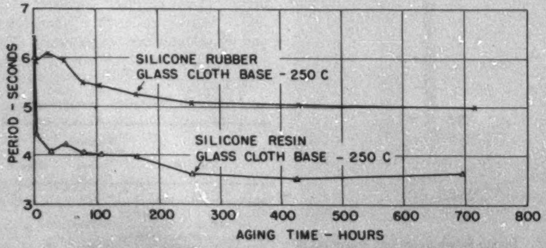


Fig. 6. Mechanical aging characteristics of silicone rubber and silicone resin on glass cloth base as measured by torsion pendulum, using pendulum no. 2

lar) at 200 C were terminated when the materials broke during test.

The consistency of results for the black varnished cloth for the different temperatures indicates that orderly chemical reactions are predominantly influencing deterioration. In contrast, the fluctuating results for polyethylene terephthalate (Mylar) are believed to be caused primarily by physical changes which are known to influence the behavior of this material in the region of 150 C. The silicone rubber impregnated glass cloth exhibits much greater stability at 250 C than does the silicone resin treated material.

Nomenclature

a = strip thickness, inches
b = strip width, inches
E = Young's modulus, pounds per inch² (lb/in²)
G = shearing modulus, lb/in²
G_0 = shearing modulus with zero longitudinal stress, lb/in²
G_0^{1} = shearing modulus of unaged strip with zero longitudinal stress, lb/in²
G_0^{11} = shearing modulus for aged strip with zero longitudinal stress, lb/in²
I = moment of inertia of mass, inch-pound-seconds²
I_1 = moment of inertia of lighter pendulum, inch-pound-seconds²
I_2 = moment of inertia of heavier pendulum, inch-pound-seconds²
L = length of suspending strip, inches
M = torque, inch-pounds
m_1 = weight of lighter pendulum, pounds
m_2 = weight of heavier pendulum, pounds
T = time of oscillation, seconds per oscillation
T_0 = period of torsion pendulum with zero weight, seconds
T_1 = period of lighter pendulum, seconds
T_2 = period of heavier pendulum, seconds
U = kinetic energy, inch-pounds
U_0 = kinetic energy for pendulum of zero weight, inch-pounds
U_1 = kinetic energy for lighter pendulum of m_1 weight, inch-pounds
U_2 = kinetic energy for heavier pendulum of m_2 weight, inch-pounds
U_0^{1} = kinetic energy of unaged strip with zero longitudinal stress, inch-pounds
U_0^{11} = kinetic energy for aged strip with zero longitudinal stress, inch-pounds
α = angular deflection, radians per inch length of strip
θ = angular deflection, radians
σ = strip longitudinal stress, lb/in²
$\psi = \alpha$ initial, radians per inch length of strip
ω = angular velocity, radians per second

Torsion Pendulum With Thin Rectangular Strips

The torque produced by the twisting of a thin rectangular strip (where ratio of width to thickness is greater than 100) is shown in equation 29 to be

$$M = \frac{a^3 b G \alpha}{3}\left(1 + \frac{E b^4 \alpha^2}{120 G a^2} + \frac{\sigma b^2}{4 G a^2}\right) \qquad (1)$$

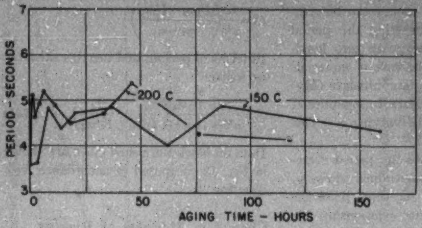


Fig. 7 (left). Mechanical aging characteristics of polyethylene terephthalate (Mylar) as measured by torsion pendulum, using pendulum no. 2

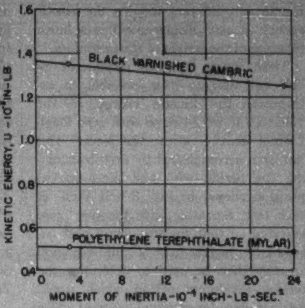


Fig. 8 (right). Kinetic energy of torsion pendulums with constant pendulum weight

This may be rearranged by collecting terms

$$M=\left(\frac{a^3bG}{3}+\frac{ab^3\sigma}{12}\right)\alpha+\left(\frac{E}{G}\right)\frac{ab^5G}{360}\alpha^3 \qquad (2)$$

$$=A\alpha+B\alpha^3 \qquad (3)$$

where

$$A=\frac{a^3bG}{3}+\frac{ab^3\sigma}{12} \qquad (4)$$

$$B=(P)\frac{ab^5G}{360} \qquad (5)$$

$$P=\frac{E}{G}$$

If a mass suspended by the long strip of rectangular cross section is twisted through the angle α and released, the mass will perform oscillatory rotation about the major axis of the strip. Neglecting the mass of the strip and air resistance, and assuming perfect elasticity and no internal friction in the strip, the equation of motion of the mass is

$$\text{(inertia)(acceleration)}+\text{restoring force} = 0 \qquad (6)$$

Since the restoring force is given by equation 3, equation 6 becomes

$$I\left(\frac{d^2L\alpha}{dt^2}\right)+A\alpha+B\alpha^3=0 \qquad (7)$$

or, for a given length

$$\frac{d^2\alpha}{dt^2}+\frac{A\alpha}{IL}+\frac{B\alpha^3}{IL}=0 \qquad (8)$$

This differential equation has the approximate solution[6]

$$\omega^2\cong\frac{A}{IL}+\frac{3}{4}\frac{B}{IL}\psi^2 \qquad (9)$$

Since

$$T=\frac{2\pi}{\omega} \qquad (10)$$

$$T^2=\frac{4\pi^2}{\frac{A}{IL}+\frac{3}{4}\frac{B}{IL}\alpha^2} \qquad (11)$$

Substituting values of A and B from equations 4 and 5

$$\frac{4\pi^2}{T^2}=\frac{4a^3bG+ab^3\sigma}{12IL}+\frac{Pab^5G}{480IL}\alpha^2 \qquad (12)$$

Solving for shearing modulus G

$$G=\frac{1920\pi^2IL-40ab^3T^2\sigma}{160T^2a^3b+T^2Pab^5\alpha^2} \qquad (13)$$

Interpretation of Results in Terms of Shearing Modulus

Certain difficulties are encountered when equation 13 is utilized to interpret the results of Figs. 5, 6, and 7 in terms of shearing modulus G. This expression for shearing modulus contains constants involving dimensions of the strip, dimensions, and weight of the oscillating mass, time of oscillation of the mass, and the quantity P. P is the ratio of Young's modulus E to the modulus of rigidity G, and is of the order of 3 in the initial unaged state for typical insulations. This ratio may change with aging since generally G and E do not change in the same proportions. However, by making α small (15 degrees) it is found that the whole term is a small percentage of the other denominator term and can be neglected with minor error.

Unfortunately the numerator of this expression is the difference of two quantities which are of the same order of magnitude for certain typical dimensions and characteristics of insulations. Thus, the computation of G using this expression is subject to considerable error. It might be expected that using a smaller b (a narrower strip) would help this since it occurs as a cubed term in equation 13. However, decreasing b causes T to increase in about the same proportion, so little is gained. Some improvement is made by making σ smaller if at the same time I is not decreased. This is accomplished by using a mass consisting of long narrow bars. A major improvement may be made in the accuracy of the shearing modulus calculation by obtaining the value of G under the condition of zero applied longitudinal stress. This may be done by considering the energy relations governing the motion of the suspended mass.

Kinetic Energy Relation

Assuming no translatory motion, the kinetic energy of the oscillating mass as it passes through zero displacement is

$$U=\frac{1}{2}I\omega^2 \qquad (14)$$

Assuming that the torsion pendulum executes simple harmonic motion (which is approximately true if displacements are small)

$$\omega^2=\left(\frac{2\pi}{T}\right)^2 \qquad (15)$$

where T = period of pendulum in seconds per oscillation. Hence

$$U=\frac{2\pi^2I}{T^2} \qquad (16)$$

The kinetic energy of the bar should be equal to that energy stored in the tape originally by displacing the bar through the angle ψ. This is verified, in part at least, by the data of Fig. 8 which show that the inertia of the mass has no significant effect on the stored energy calculated from equation 16. This may be confusing, since equation 16 indicates U proportional to I. Note, however, that U is also inversely proportional to T^2 which changes in such a manner with I that the over-all effects cancel out.

It is apparent even from the ideal theory of equation 1 that the energy stored in the test strip will be affected by the longitudinal stress due to the weight of the bar. Test results illustrated in Figs. 9 and 10 show that for torsion pendulums of equal moments of inertia and different weights the kinetic energy at zero displacement is a linear function of the weight of the pendulum.

Thus, using torsion pendulums of different weights it is possible to extrapolate the kinetic energy to the value U_0 it would have with zero longitudinal stress σ on the test piece. This has been accomplished by first determining the period with test piece in the position shown in Fig. 3 and then, at each test, repeating with the test piece inverted so that the top clamp now becomes the inertia bar. U_0 and equation 16 may then be used to determine the torsion pendulum period which would occur if no longitudinal stresses were applied to the test strip. This relation is of great importance and utility since with zero stress σ the term involving σ in equation 13 becomes zero. Thus the calculation of G is more accurate because it no longer involves the difference of two quantities that may be nearly alike. Also, it allows investigation of materials which are so nonisotropic that equation 13 does not give a proper allowance for the influence of longitudinal stresses. For example, the apparent shearing modulus of black varnished cloth as calculated from equation 13 is affected by changing σ; see Fig. 11. This deviation from ideal theory can be compensated for by correcting the period to correspond to that with zero longitudinal stress. In the case of materials such as polyethylene terephthalate (Mylar) equation 13 appears to compute G consistently even with varying σ. This is shown in Table III.

In order to calculate the period corresponding to zero longitudinal stress σ, it is only necessary, for those materials which exhibit a linear relationship between U and σ (as illustrated in Figs. 9 and 10), to compute the y intercept on the curve of weight versus kinetic energy. Then

$$U_0=\frac{U_1m_2-U_2m_1}{m_2-m_1} \tag{17}$$

With σ equal to zero and for small displacement equation 13 becomes

$$G_0=\frac{12\pi^2IL}{T_0^2a^3b} \tag{18}$$

Substituting the value of T from equation 16 with appropriate subscripts and simplifying

$$G_0=\frac{6LU_0}{a^3b} \tag{19}$$

Utilizing equation 17

$$G_0=\frac{6L}{a^3b}\left(\frac{U_1m_2-U_2m_1}{m_2-m_1}\right) \tag{20}$$

In terms of periods and inertias this may be written

$$G_0=\frac{12L\pi^2m_1m_2}{a^3b(m_2-m_1)}\left(\frac{I_1}{T_1^2m_1}-\frac{I_2}{T_2^2m_2}\right) \tag{21}$$

Data on black varnished cloth are shown in Fig. 12 computed in accordance with equation 21.

Change in Modulus of Rigidity With Insulation Aging

Actually of more concern is the change in G rather than its actual value. From equation 19

$$G_0^{1}=\frac{6LU_0^{1}}{a^3b} \tag{22}$$

Also

$$G_0^{11}=\frac{6LU_0^{11}}{a^3b} \tag{23}$$

Then dividing equation 23 by 22

$$\frac{G_0^{11}}{G_0^{1}}=\frac{U_0^{11}}{U_0^{1}} \tag{24}$$

Further Considerations

In some cases there appears to be an effect on pendulum period due to a

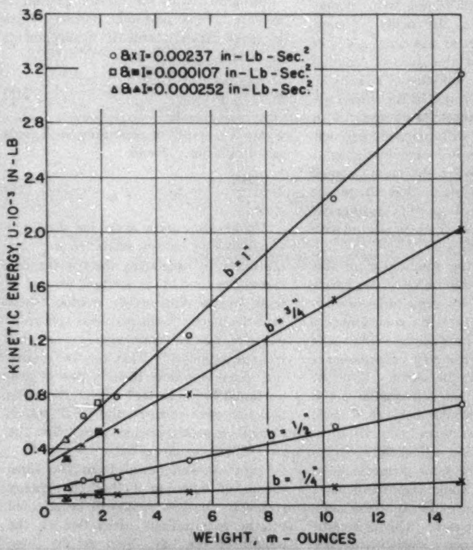


Fig. 9. Kinetic energy of torsion pendulums utilizing black varnished cloth with varying pendulum weight

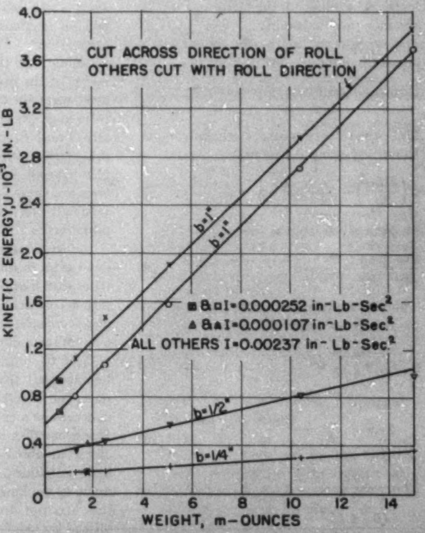
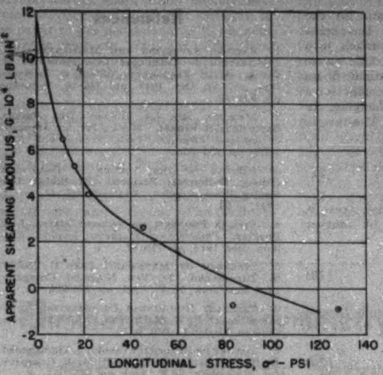


Fig. 10. Kinetic energy of torsion pendulums utilizing polyethylene terephthalate (Mylar) with varying pendulum weight



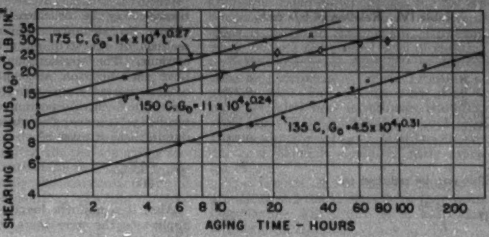


Fig. 11 (left). Effect of varying longitudinal stress with pendulum inertia kept constant

Fig. 12 (above). Shearing modulus for black varnished cloth as affected by aging, calculated from torsion pendulum data

damage of the edge of the material when cutting the sample to width. For example, in the case of black varnished cloth, Fig. 9, the ordinate U_0 is not directly proportional to b as specified in equation 19. However, if a constant amount is deducted from the quantity b to allow for this damage, reasonably consistent results for various widths are obtained. Some other materials such as polyethylene terephthalate (Mylar) show little or no edge effect; see Table III. It thus appears wise to use tapes not less than 1 inch wide to minimize this effect.

In some materials there is an effect on pendulum period due to the direction of cutting as referred to the roll from which the sample is taken. This goes back to the method of manufacture, fiber orientation, etc. An illustration is given in Fig. 10 where samples were cut at right angles to each other. It is thus necessary to be consistent in sampling or else investigate separately the variations encountered.

There may be some effect of ambient conditions on the pendulum period. This has not been fully investigated but it may be desirable to cool for a time in excess of 1 hour and test samples in a room with controlled temperature and humidity.

Deterioration Rates

The orderly changes in the shearing modulus as a function of aging time indicate that this test procedure may prove useful in the study of mechanical deterioration rate phenomena. Data on black varnished cloth at 135, 150, and 175 C (see Table II) show that over a significant portion of insulation life the material exhibits a linear relation between shearing modulus, as calculated by the use of equation 21, and a fractional power of the aging time, as in Fig. 12. The highest values of G_0 shown were obtained with the material in a very brittle and therefore failing condition. It should be noted that the rate of change of G_0 with time is greater for higher temperatures.

The initial variation in torsion pendulum period and consequently in G_0, indicates the desirability of preconditioning at a temperature of the order of 100 C for several hours to give consistent data at initial conditions. It may also be important to be sure of stable measuring conditions by cooling a sufficient time under controlled temperature and humidity.

Table III. Shearing Modulus for Polyethylene Terephthalate

Nominal Thickness 8 Mils

Moment of Inertia, I(10^{-4} Inch-Pound-Seconds^2)	Weight, Ounces	σ, Lb/In^2	Width, Inches	Period T, Seconds	Modulus G(10^5 Lb/In^2)
2.524	0.6094	4.88	1	1.56	2.54
1.074	1.828	14.65	1	0.94	2.66
23.8	5.06	39.6	1	3.76	2.73
2.524	0.6094	9.77	1/2	2.46	2.38
1.075	1.828	29.29	1/2	1.54	2.39
23.8	5.06	79.2	1/2	6.68	2.16
2.524	0.6094	19.53	1/4	3.44	2.50
1.075	1.828	58.59	1/4	2.18	2.56
23.8	5.06	162.18	1/4	9.80	2.46
3.065	4.97	159.29	1/4	3.44	2.69

It is recognized that additional work is necessary, in order to investigate more fully extraneous effects, before data such as in Table II can be profitably analyzed to show the true interrelation of time and temperature on shearing modulus. The data are presented in order to show the basic methods of analysis and computation rather than final results, in the hope that other investigators will be encouraged to attack the insulation aging problem by this new and intriguing technique.

Conclusions

1. A new method has been developed for determining the shearing modulus of flexible sheet materials.

2. This torsion pendulum technique provides a nondestructive method for determining significant mechanical aging characteristics of flexible sheet materials.

3. This procedure should be useful in establishing the deterioration rates of an important mechanical property of various sheet materials.

4. This method offers a means of investigating the co-ordination of mechanical and electrical deterioration rates of sheet insulation materials.

Appendix I. Torque for Thin Rectangular Strips

The equation relating torque to angle of twist for a rectangular strip in such a material as steel is given by Timoshenko[7] as

$$M=\frac{G\theta a^3 b}{3L}\left[1-\frac{192\,a}{\pi^5 b}\sum_{n=1,3,5,\ldots}^{\infty}\frac{\tanh}{n^5}\left(\frac{n\pi b}{2a}\right)\right] \quad (25)$$

Table IV. Relation of b/a to k_1

b/a	k_1
1.0	0.1406
2.0	0.229
5.0	0.291
10.0	0.312
100.0	0.331
∞	0.333

This can be reduced to

$$M=\frac{k_1G\theta a^3b}{L} \tag{26}$$

where k_1 is a numerical factor depending on the ratio b/a. See Table IV.

Equation 26 is very accurate for such material as steel in usual shapes in which the maximum shearing strain during torsion is very small. However, for such a material as rubber and electrical insulating materials, or even very thin strips of steel, the maximum shearing strains during torsion may be considerable and so the change in distance between the cross sections of the strip must be taken into account. The stresses due to this cause have the direction of longitudinal fibers of the twisted strip and are compressive near the central portion, changing to tensile stresses near the edges. The torque due to these stresses is[5]

$$M=\frac{Eab^5}{360}\theta^3 \tag{27}$$

Also, if a uniform longitudinal stress is applied to the strip there will be another added torque[4,5]

$$M=\frac{\sigma\theta ab^3}{12} \tag{28}$$

Combining equations 25 through 28 and assuming $b/a>100$

$$M=\frac{a^3bG\theta}{3L}\left(1+\frac{Eb^4\theta^2}{120Ga^2}+\frac{\sigma b^2}{4Ga^2}\right) \tag{29}$$

References

1. Factors Affecting the Mechanical Deterioration of Cellulose Insulation, **F. M. Clark.** *AIEE Transactions* (*Electrical Engineering*), vol. 61, Oct. 1942, pp. 742–49, discussion pp. 991–92.
2. Memoire Sur La Torsion Des Prismes, **Barre deSaint-Venant.** *Mem. Acad. Sci. Memoires des Savants Etrangers*, Paris, France, vol. XIV, 1855, pp. 233–560.
3. History of the Theory of Elasticity (book), **Todhunter, Pearson.** Cambridge University Press, London, England, vol. II, pt. 1, 1893, p. 23.
4. Bifilar Property of Twisted Strips, **J. C. Buckley.** *Philosophical Magazine*, London, England, Dec. 1914, pp. 778–87.
5. Strength of Materials, Part II (book), **S. Timoshenko.** D. Van Nostrand Company, Inc., New York, N. Y., 1941, pp. 298–304.
6. Ordinary Non-Linear Differential Equations (book), **N. W. McLachlan.** University Press, London, England, 1950, p. 33.
7. Theory of Elasticity (book), **S. Timoshenko, J. N. Goodier.** McGraw-Hill Book Company, Inc., New York, N. Y., 1951, pp. 275–78.

◆

Discussion

G. A. Cypher (General Electric Company, West Lynn, Mass.): The authors have presented an interesting and valuable method for following the thermal degradation of a mechanical property of insulating sheet materials. The nondestructive nature of the test and its applicability to nonisotropic materials make it especially useful. The physical condition of materials like impregnated glass fabric after heat aging has been difficult to assess because of the strength and stability of the inorganic fabric. The condition of the organic or silicone treating material, most important from the electrical point of view, may be determined by this torsion pendulum technique.

The electrical properties of some of the materials mentioned in this paper have been studied as a function of time and temperature. The authors' comments would be appreciated on any possible relationship which may exist between some electrical characteristic as a function of aging and these values of shearing modulus as a function of aging.

R. C. Bartlett (Natvar Corporation, Woodbridge, N. J.): The authors deserve a word of commendation, since their testing procedure is uniquely different from any used so far for measuring thermal stability of insulation.

Although most of the data presented indicates that this measurement of embrittlement of insulation by a shearing modulus test is adequately quantitative, the ethylene-glycol terephthalate polyester film (Mylar) gives test values that are outstandingly more qualitative in nature. It is suggested that these more erratic results obtained with this polyester film could be explained by the relaxation of orientation in this film. Heating at 150 and 200 C might cause dimensional changes that would affect the period of oscillation and calculated shearing modulus. Although little information on thermal relaxation and accompanying dimensional changes of such polyester film has been published, an investigation of this might shed some light on the erratic results obtained.

It is well known that thickness of insulation will change during heat aging. Some data of this kind have been given previously.[1] Since the equation given to calculate shearing modulus includes a thickness dimension of the test specimen, it would aid evaluation of this test procedure to know whether progressive change in specimen thickness was considered in the data presented. This point has not been made clear in this paper, and it might be worth considering thickness change in insulation if the authors have not already done so.

The authors have given an opinion on significance of such shearing modulus data when they say, "The silicone rubber-impregnated glass cloth exhibits much greater stability at 250 C than does the silicone resin treated material." It is true that initial changes in the oscillation period of the silicone-varnished glass cloth are much greater than for the silicone rubber-coated fabric. No doubt the authors have based their statement on this fact. However, it is worth noting that after 200 hours of accelerated aging slope of both the silicone varnish and silicone rubber insulation in the oscillation period versus aging time curves are nearly the same. Since the longer term data indicate nearly similar decay rates, I should like to suggest that the thermal stability of both these insulations is substantially the same.

I trust these comments will aid the understanding and use of the shearing modulus test method and data presented by the authors.

Reference

1. Thermal Stability of Insulating Fabrics, R. C. Bartlett. *Publication No. 161*, American Society for Testing Materials, Philadelphia, Pa.

L. C. Whitman and A. L. Scheideler: Dr. Cypher has brought out a very important point in the utility of the torsion pendulum technique in assessing mechanical characteristics of materials having a base material like glass fabric. Such materials maintain a major portion of their mechanical strength in tension irrespective of aging and the consequent deterioration of the impregnant. Thus tensile tests do not provide a good criterion of aging for such materials. The shearing modulus as tested by the torsion pendulum technique, however, does test the mechanical condition of the impregnant. It is this impregnant that is important as far as electrical characteristics of the insulating material are concerned. Thus, the shearing modulus as measured by the torsion pendulum can logically be expected to correlate with electrical characteristics. The authors are, however, accumulating additional data on a variety of materials and expect to report on this information at a later date.

Mr. Bartlett suggests that the somewhat erratic results obtained using polyethylene terephthalate film (Mylar) are due to relaxation of orientation in this film, together with dimensional changes. There were noticeable changes in length during aging, these changes also being of an irregular nature. It should be emphasized that the pendulum periods were at all times repeatable within an accuracy of about 2 per cent so that the large erratic changes shown are not due to to experimental variations but reflect actual changes in characteristics, as Mr. Bartlett suggests.

The initial thickness measurement was used in the calculation of shearing modulus in the data shown. Using actual thickness at each test is of course more nearly correct and in the case of large thickness changes will alter results, especially since thickness as a cubed term. Inclusion of this effect will tend to increase the calculated shearing moduli as aging time increases and even probably to a greater extent at the higher temperatures. Thus decreases in thickness of the order of 10 per cent will increase the calculated shearing moduli in the order of 27%. In the case of black varnished cambric at 150 C the change in shearing modulus for 40 hours would be increased from the value of 265 per cent as shown in the paper to 340 per cent for a 10 per cent change in thickness. It is not evident how this will affect the shape of the aging curve. Inclusion of this effect, however, does make the shearing modulus criterion even more sensitive. We are indebted to Mr. Bartlett in pointing out this considerable effect.

In regard to the comparison of silicone resin and silicone rubber on a glass cloth base, it is evident that while these particular test samples initially had approximately the same period that their initial shearing moduli were quite different. The thickness of the silicone resin-glass sample was one-half that of the silicone rubber-glass sample (see Table I). Hence the initial shearing modulus of the silicone resin-glass would be of the order of eight times that of the silicone rubber-glass sample. After 700 hours the periods were in the ratio of 3.5 to 5.0. Since shearing modulus under zero longitudinal stress conditions is inversely proportional to period squared, this 8-to-1 ratio would be reduced by a factor of the square of the ratio of the periods giving an over-all result of about 3.9 to 1, so that both initially and finally this particular silicone rubber-glass combination had a considerable edge on the silicone resin-glass as far as flexibility is concerned. It is true that, after some rapid initial changes and for the temperatures and times shown, the rate of aging appears to be similar. However, when we note the general level of shearing modulus (and hence flexibility or brittleness) as approximated in the foregoing it appears that the silicone resin-glass is probably fairly near its ultimate condition while the silicone-rubber-glass would seem to have a long way to go. Only longer tests on these materials would definitely establish this.

The interest shown in this method is gratifying and the authors hope that other investigators will utilize it in evaluating flexible sheet materials.

Elimination of Asymmetry Zero-Drift Errors in Magnetic Servo Amplifiers

W. A. GEYGER
MEMBER AIEE

Synopsis: When applying magnetic amplifiers to position-indicating synchro-transmission systems, achievement of an extremely small zero drift is of paramount importance. It is well known that zero-drift errors caused by actual deviations from perfect symmetry of the saturable-reactor and rectifier components of magnetic-amplifier push-pull circuits may be reduced to certain specified values by using proper testing procedures for grading and matching of these components. However, limitation of such errors to sufficiently small values still represents a difficult problem in the design and production of high-performance magnetic servo amplifiers. Application of 2-phase induction motors in connection with special circuit arrangements offers the possible elimination of asymmetry zero-drift errors by supplying the magnetic-amplifier circuit from a magnetic frequency multiplier producing an exact multiple of the frequency of the a-c line feeding directly both the line phase of the motor and the synchro generator of the error-detecting system. In this case, even with nonsymmetrical characteristics of the saturable-reactor and rectifier components, operation of the motor under no-signal conditions will be inherently independent of changes in magnitude and frequency of the a-c line voltage, changes in ambient temperature, etc., with regard to the magnetic amplifier. Application of a higher frequency carrier for supplying the magnetic servo amplifier offers other advantages: its operation becomes more efficient and its speed of response is correspondingly increased. Actually, the possibility of using smaller sized saturable-reactor elements may compensate somewhat for the necessity of additional components of the magnetic frequency multiplier.

WHEN applying magnetic servo amplifiers to position-indicating synchro-transmission systems, the amplifier is supposed to have no output, if its input voltage is zero as a result of exact coincidence of alignment between the control-transformer rotor and the synchro generator rotor, and it is required that fulfillment of this condition be independent of changes in magnitude and frequency of the a-c line voltage, changes in ambient temperature, etc.

The conventional way of applying magnetic amplifiers of the balance detector type[1] to position-indicating synchro-transmission systems consists in operating the magnetic-amplifier push-pull circuit directly from the same 60-cycle or 400-cycle power source from which the synchro system is operated, connecting the input of the amplifier to the synchro-control transformer, and controlling a 2-phase induction-type reversible motor, geared to the control transformer, with the output of the amplifier. The control transformer furnishes an a-c signal voltage, the magnitude of which is an indication of the size of the angular error between control-transformer rotor and synchro-generator rotor, while its polarity or phase is an indication of the direction of the error. Thus, the synchro control transformer performing the function of the differential element of the error-detecting system serves to compare actual angular positions of the input and putput members of the system.

Since the magnetic amplifier is directly connected with the a-c line feeding the line phase of the motor, any undesired output caused by zero-drift errors of the amplifier itself will cause the motor to operate exactly in the same way as does the desired signal-proportional output. Thus, the 2-phase motor, when operated in this way, is unable to discriminate between error-signal output and zero-drift output of the magnetic amplifier.

The Asymmetry Zero-Drift Problem

Magnetic-amplifier push-pull circuits,[2] as used with synchro systems, represent symmetrical a-c bridge networks containing saturable core reactors and dry-disk rectifiers. With perfect symmetry of the saturable-reactor and rectifier components, the aforementioned condition, i.e., the balance condition of the a-c bridge network, will be fulfilled, even when changes in power supply voltage frequency, and ambient temperature occur.

In practice, however, some deviations from perfect symmetry will always exist, because respective characteristics of the saturable-reactor and rectifier elements will never be perfectly identical. Therefore, with zero input, a certain comparatively small output corresponding to the actual amount of asymmetry will be produced. This output, the "asymmetry zero-drift error" of the magnetic amplifier, should be as small as possible.

It is well known that zero-drift errors caused by actual deviations from perfect symmetry of the saturable-reactor and rectifier components may be reduced to certain specified values by using proper testing procedures for the grading and matching of these components. However, the limitation of such errors to sufficiently small values still represents a difficult problem in the design and pro-

---

Paper 55-59, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted September 22, 1954; made available for printing December 15, 1954.

W. A. GEYGER is with the United States Naval Ordnance Laboratory, Silver Spring, Md.

duction of high-performance magnetic servo amplifiers.

Elimination of Asymmetry Zero-Drift Errors

An obvious way of eliminating zero-drift errors in magnetic servo amplifiers consists in operating the magnetic-amplifier push-pull circuit as a real a-c amplifier from a separate, higher frequency power source and feeding both the line phase of the 2-phase motor and the synchro generator of the error-detecting system from the 60-cycle or 400-cycle a-c line. This method, of course, makes it necessary to use an additional higher frequency power source, which is generally not available. However, it may be possible in some cases, e.g., in feeding both the line phase of the motor and the synchro generator from the 60-cycle line, whereas the magnetic amplifier will be operated from a 400-cycle power source.

Application of special circuit arrangements, as described in this paper, offers the possibility of eliminating asymmetry zero-drift errors by supplying the magnetic-amplifier push-pull circuit from a magnetic frequency multiplier producing an exact multiple of the frequency of the 60-cycle or 400-cycle a-c line, feeding directly both the line phase of the motor and the synchro generator.

When applying this special method of magnetic-amplifier operation, the output-circuit elements of the amplifier will carry alternating and/or unidirectional currents, the frequency of which is a multiple of the fundamental frequency of the a-c line voltage. For example, when employing a 400/1,200-cycle magnetic frequency tripler for this purpose, a carrier frequency of 1,200 or 2,400 cycles per second (cps) is applied to the output-circuit elements of the magnetic amplifier, while the line phase of the motor and the synchro generator are directly supplied from the 400-cycle line. Therefore, the 400-cycle signal voltage of the synchro control transformer is used for modulating the 1,200-cycle output current of the magnetic amplifier which does feed the control phase of the motor. Thus, in this case, the magnetic amplifier operates as a real a-c amplifier with a carrier frequency $f_c = 1{,}200$ cps, a modulating frequency $f_m = 400$ cps (which is in turn modulated at the signal frequency), and the frequency ratio $f_c/f_m = 3$. This is in contrast to the conventional techniques using the frequency ratio $f_c/f_m = 1$.

It is interesting to note that in this special case of a nulling device it has proved possible to apply the unusually low frequency ratio $f_c/f_m = 3$. This is in contrast to conventional modulation applications where f_c/f_m should be at least 7 to 10.

When applying this special combination of magnetic servo amplifier, 2-phase motor, and magnetic frequency tripler, any undesired output caused by zero-drift errors of the amplifier itself consists of a 1,200-cycle or 2,400-cycle, and a direct current flowing in the control phase of the motor. This current, of course, will not operate the motor (provided that any fundamental component in the carrier does not affect the output of the amplifier), the line phase of which is carrying a 400-cycle current. However, the desired signal-proportional 400-cycle output current of the amplifier will efficiently operate the motor in the usual way. Thus, in this case, the 2-phase motor is able to discriminate between amplifier output values due to application of 400-cycle error voltage from the synchro control transformer and those 1,200-cycle or 2,400-cycle and direct current amplifier output values which may be caused by zero-drift errors produced by actual deviations from perfect symmetry of the saturable-reactor and rectifier components of the magnetic amplifier.

In this way, asymmetry zero-drift errors will be eliminated and, even with nonsymmetrical characteristics of these components, operation of the motor under no-signal conditions will be inherently independent of changes in magnitude and frequency of the a-c line voltage, changes in ambient temperature, etc., with regard to the magnetic amplifier.

Also, of course, application of a higher frequency carrier for supplying the magnetic servo amplifier offers other important advantages in so far as the operation of this amplifier becomes more efficient and as its speed of response is correspondingly increased. Actually, the possibility of using smaller sized saturable-reactor elements may to a certain extent offer some compensation for the necessity of providing the additional components of the magnetic frequency multiplier. Furthermore, it may be stressed that magnetic amplifiers and magnetic frequency multipliers have almost unlimited life. Thus, it is possible to seal both these components hermetically in a single compact case.

At this point it is to be noted that the use of magnetic frequency multipliers in connection with magnetic amplifiers to decrease the size of the saturable-reactor elements and to increase the speed of response was suggested many years ago.[3] This paper, however, deals particularly with the application of special circuit arrangements with magnetic servo amplifier, magnetic frequency multiplier, and 2-phase motor to position-indicating synchro-transmission systems to eliminate asymmetry zero-drift errors of the magnetic amplifier.

The numerous possibilities for applying this special method of magnetic-amplifier operation may be illustrated by a typical example of a practical circuit arrangement which has proved that it can give excellent results when used in connection with instrument-type servos.

Position-Indicating Synchro-Transmission System with 2-Phase Motor, 2-Stage Half-Wave Magnetic Servo Amplifier, and Magnetic Frequency Tripler

The preference expressed for half-wave types of multistage magnetic servo amplifier circuits[4,5] spurred development work on such circuits to be used in connection with various single-phase and 3-phase types of magnetic frequency multiplier circuits, both for line frequencies of 60 and 400 cps.

Fig. 1 shows schematically such a system that will be operated, e.g., from an a-c line with $E_P = 115$ volts, $f_P = 400$ cps, in connection with a 400/1,200-cycle, single-phase, or 3-phase tripler circuit.

In accordance with the basic principle of eliminating asymmetry zero-drift errors, the magnetic-amplifier circuit is connected with the 1,200-cycle output of the magnetic frequency tripler, the 400-cycle input of which is supplied from the 115 volts, 400-cps, a-c line feeding directly both the line phase W_P (with series capacitor C_P) of the 2-phase motor, and the rotor winding R_1R_2 of the synchro generator. The stator windings $S_1S_2S_3$ of the synchro generator are connected with those of the synchro control transformer, the rotor winding R_1R_2 of which is connected with the input of the amplifier. The control phase W_L of the 2-phase motor, geared to the control transformer, is supplied from the output stage of the amplifier with phase-reversible output current I_L.

Output stage and input stage of the half-wave push-pull magnetic-amplifier circuit, preferably having saturable-reactor elements with Orthonol[6–8] tape cores, are coupled by means of differential control windings, N_C' and N_C''. The graphic symbols of Fig. 1, which can be used in a general way for illustrating the actual mode of operation of magnetic-amplifier circuits,[3] show that the black rectifiers D_A' and D_A'' are conducting simultaneously during the first half-cycle period of 1,200-cycle supply voltage, E_M (black polarities), while the white rectifiers D_D' and D_D'' conduct simultaneously during the second half-cycle period of E_M (white polarities). Thus, the output-stage circuit is supplied with unidirectional currents I_A' and I_A'' (black half-cycle pulses), whereas the input-stage circuit carries the unidirectional currents I_D' and I_D'' (white half-cycle pulses), so that the cores of the output-stage circuit are controlled differentially by means of input-stage currents I_D' and I_D'' during the "reset" half-cycle periods.

It is to be noted that the frequency of the a-c components of unidirectional currents I_A' and I_A'', and I_D' and I_D'', is 1,200 cps. Therefore, any 1,200-cycle output caused by asymmetry zero-drift errors of the magnetic amplifier itself will not operate the 2-phase motor, the line phase W_P of which is carrying 400-cycle current ${}_P$. However, when a 400-cycle error-signal voltage E_S is applied to the control circuit $N_S'N_S''R_S$ of the input stage, then currents I_A' and I_A'', flowing through the differential primary windings N_1' and N_1'' of output transformer T_L, will be modulated with 400 cps, and a properly phased 400-cycle output current I_L, flowing through the control phase W_L, will operate the motor in the usual way. Nearly sinusoidal wave shape of output current I_L is obtained as a result of the demodulating effect of a properly rated capacitor C_L connected across the secondary winding N_2 of output transformer T_L.

When the angular position α_G of the synchro generator is exactly coincident with the angular position α_{CT} of the synchro control transformer, there will be no error-signal voltage E_S, output current I_L will be zero, and the motor will remain at rest. However, if the controlling input member M_α, mechanically coupled with the synchro generator, is displaced from its original position, then an error-signal voltage E_S will energize the amplifier, the output current I_L of which will cause the motor to operate in the proper direction to bring the rotor of the synchro control transformer in alignment with the rotor of the synchro generator.

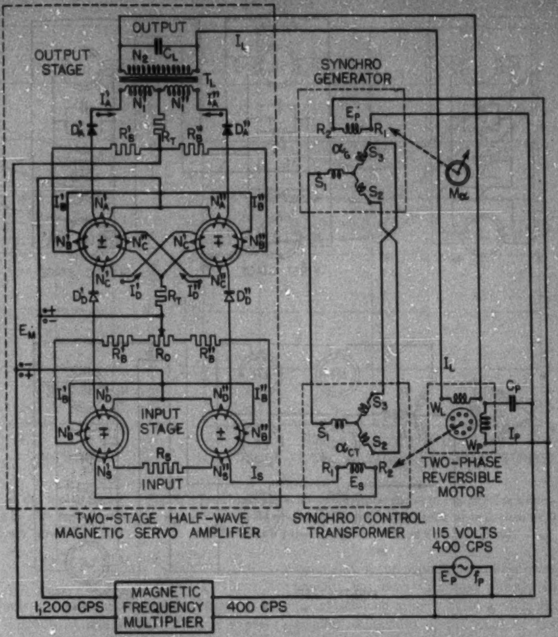


Fig. 1. Position-indicating synchro-transmission system, the 2-phase motor of which is operated by special combination of 2-stage half-wave magnetic servo amplifier and magnetic frequency multiplier to eliminate asymmetry zero-drift errors of amplifier

Magnetic Frequency Multiplier Circuits

The basic principle of magnetic frequency multipliers, sometimes called "static frequency transformers" or "harmonic generators," consists in producing higher harmonics of the power-supply frequency by saturation of magnetic cores, and eliminating the fundamental component and other undesired components by means of bridge and/or filter circuits.

Various single-phase and polyphase types of magnetic frequency multipliers containing saturable reactors, linear reactors, filter circuits and also, in some cases, rectifiers, have been devised by early investigators in the field of radio engineering.[9] Such circuit arrangements with the use of either "sinusoidal excitation" or "shock excitation" were applied many years ago for supplying ink-recording frequency meters[10] and a-c bridge and potentiometer networks.[11] Some magnetic frequency multiplier circuits, proved suitable for supplying magnetic amplifiers, have been described recently.[12–14]

It has been stressed that the fundamental component of the output voltage of the magnetic frequency multiplier should be sufficiently small not to affect the output of the magnetic amplifier. Results of measurements, to be described, have shown that fundamental component distortions in the order of 5 to 10 per cent (%) have no appreciable effect with regard to elimination of asymmetry zero-drift errors caused by very large deviations from perfect symmetry of the saturable-reactor and rectifier components of the magnetic-amplifier circuit of Fig. 1.

Single-Phase Circuits

Fig. 2(A) shows a circuit diagram of a simple single-phase bridge-type[15] magnetic frequency tripler representing a modification of the first transformer-type arrangement[16]. Each of the windings of the saturable reactor *SR* and linear (air-gap) reactor *LR* is divided into two equally rated parts which are insulated from each other; and a properly tuned tank circuit L_TC_T is connected across the load. This tripler circuit, when carefully designed, operates with an efficiency

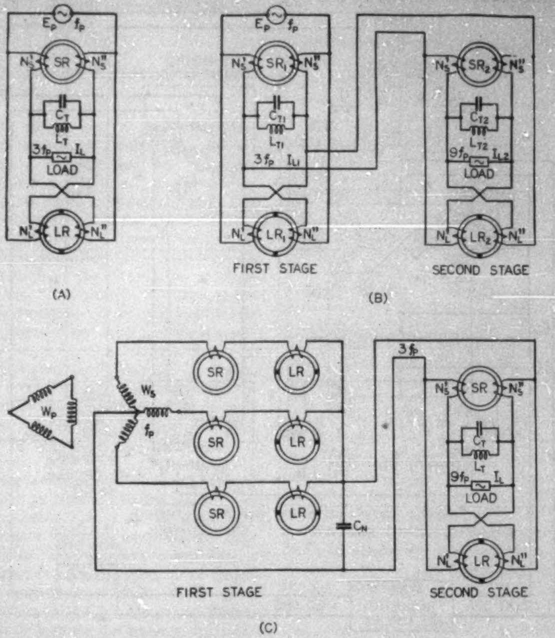


Fig. 2. Magnetic frequency multipliers to be used in connection with the arrangement of Fig. 1

A—Single-phase bridge-type frequency tripler
B—Two-stage bridge-type frequency multiplier
C—Two-stage frequency multiplier, which is a combination of 3-phase input-stage tripler and single-phase output-stage tripler

of about 70 to 80% and will provide a nearly sinusoidal output voltage, the fundamental component distortion being in the order of 5 to 10%.

It is also possible to employ a 2-stage arrangement,[13] Fig. 2(B), each stage consisting of a bridge network with a saturable reactor, a linear reactor, and a properly tuned tank circuit. This arrangement makes it possible, e.g., to convert an input power of 60 cps into an output power of $3\times3\times60=540$ cps with satisfactory efficiency (about 50 to 60%).

Three-Phase Circuits

Another possibility consists in applying the conventional method of producing third-harmonic frequency power by means of three identical saturable-core transformers[18] with wye-connected primary windings and open-Δ-connected secondary windings; and a properly rated capacitor in the secondary circuit may be used to improve the performance of this arrangement.[19]

Fig. 2(C) shows a circuit diagram of a 2-stage 60/540-cycle or 400/3,600-cycle magnetic frequency multiplier, which is a combination of a 3-phase input-stage tripler and a 1-phase output-stage tripler. The output-voltage wave shape of the 3-phase tripler circuit of Fig. 2(C), which may, of course, also be used for single-stage 60/180-cycle or 400/1,200-cycle applications, is purely sinusoidal, provided that the 3-phase loads are perfectly balanced. In this case, total harmonic distortion does not exceed 0.5 to 1%.

Such arrangements,[20] when supplied from a 3-phase, 4-wire power source, contain three wye-connected saturable reactors, *SR*, preferably with series-connected linear reactors *LR*, and a properly rated capacitor C_N; Fig. 2(C), or tank circuit across the load.

In the case of employing a 3-phase, 3-wire power source, a 3-phase transformer with Δ-connected primary windings W_P and Y-connected secondary windings, W_S must be provided, Fig. 2(C), for supplying this type of tripler circuit.

Measurements on Single-Phase Bridge-Type Magnetic Frequency Tripler

Actual performance characteristics of a single-phase, bridge-type, 400/1,200-cycle tripler circuit, Fig. 2(A), may be illustrated by considering results of various measurements on a typical arrangement, designed for a 1,200-cycle maximum output of 50 watts, that is operated from a 115-volt, 400-cycle line voltage in connection with a 2-stage half-wave magnetic amplifier, as shown in Fig. 1.

Dimensions of Components

The saturable reactor *SR* consisted of a spirally wound 2-mil Orthonol[6,7] tape core having the following dimensions: Outer diameter $1^3/_4$ inches; inner diameter $1^1/_4$ inches; tape width 1.0 inch. This core had two multilayer toroidal windings of 135 turns each, of enamel-insulated no. 20 copper wire (Brown and Sharpe gauge).

The linear reactor *LR* consisted of a double-size Permalloy dust core having the following dimensions: Outer diameter 1.85 inches; inner diameter 0.95 inch; total height 1.4 inches. This double core had two multilayer toroidal windings of 160 turns each, of enamel-insulated no. 20 copper wire (Brown and Sharpe gauge).

The linear reactor L_T of the tank circuit consisted also of a double-size Permalloy dust core the dimensions of which were the same as those of the linear reactor *LR*. This double core, however, had two parallel-connected multilayer toroidal windings of 80 turns each, of enamel-insulated no. 18 copper wire (Brown and Sharpe gauge).

Performance Characteristics

Figs. 3 and 4 show 1,200-cycle output voltage $E_L=I_L\times R_L$, and total input current I_P, as a function of 400-cycle line voltage E_P, with capacitance of capacitor C_T of the tank circuit as a parameter (resistive load $R_L=300$ ohms). Similar results were obtained when this resistive load was replaced by the magnetic-amplifier circuit of Fig. 1. Evidently, the regulation of magnitude and frequency of line voltage E_P should be limited to about ±5%.

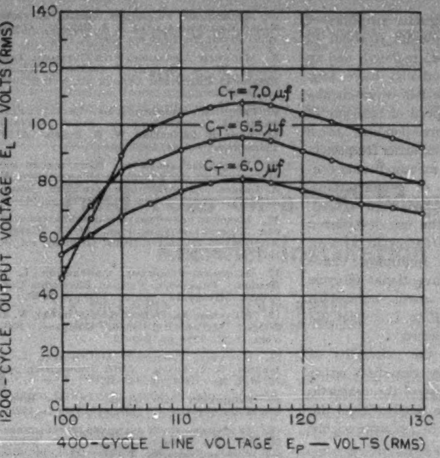


Fig. 3. 1,200-cycle output voltage $E_L = I_L \times R_L'$ of the tripler circuit of Fig. 2(A) as a function of 400-cycle line voltage E_P with capacitance of tank-circuit capacitor C_T as a parameter (resistive load $R_L = 300$ ohms)

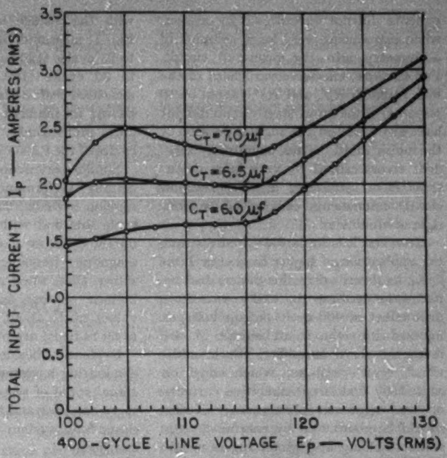


Fig. 4. Total input current I_P of the tripler circuit of Fig. 2(A) as a function of 400-cycle line voltage E_P with capacitance of tank-circuit capacitor C_T as a parameter (resistive load $R_L = 300$ ohms)

MEASUREMENT OF FUNDAMENTAL-COMPONENT DISTORTION

Actual magnitude of the fundamental 400-cycle component E_{400}, of 1,200-cycle output voltage E_L of the tripler circuit, Fig. 2(A), has been measured by applying output voltage E_L to the voltage circuit of a single-phase wattmeter, the current circuit of which was separately excited by means of a sinusoidal 400-cycle auxiliary current I_{400}, as illustrated in Fig. 5.

If the phase shifter is properly adjusted so that the wattmeter shows maximum deflection

$$P_{\max} = E_{400} \times I_{400} \times \cos\theta \qquad (1)$$

$$\theta = \angle(E_{400}, I_{400}) = 0 \qquad (2)$$

then the fundamental component of 1,200-cycle output voltage E_L is given by

$$E_{400} = P_{\max}/I_{400} \qquad (3)$$

and

$$D_{400} = \frac{E_{400}}{E_L} \times 100\% \qquad (4)$$

represents the percentage fundamental component distortion of output voltage E_L of the magnetic frequency multiplier.

This procedure is based upon the simple and convenient method for analyzing the wave shape of alternating currents by means of an electrodynamometer.[21,22]

Fig. 6 shows actual fundamental component distortion of load voltage E_L across the magnetic-amplifier circuit as a function of line voltage E_P. It is evident that magnitude of the fundamental 400-cycle component is about 6 to 8 per cent of the 1,200-cycle load voltage E_L.

Measurements on Closed-Loop Servo System

MEASUREMENT OF ASYMMETRY ZERO-DRIFT ERROR

Actual magnitude of zero-drift errors caused by very large deviations from symmetry of the magnetic-amplifier components has been measured directly in the closed-loop system by observing the angular deflections of the shaft of the 2-phase motor by means of a small mirror and a light beam, as conventionally used in mirror-type galvanometers. This arrangement has been carefully calibrated by injecting an additional 400-cycle signal voltage of 0.50 volt (corresponding to a drift rate of 0.50 degree) into the synchro control transformer circuit. It was found that with the measured fundamental component distortions of about 6 to 8 per cent of 1,200-cycle voltage across the magnetic-amplifier circuit, actual magnitude of zero-drift errors caused by asymmetry of the push-pull circuit com-

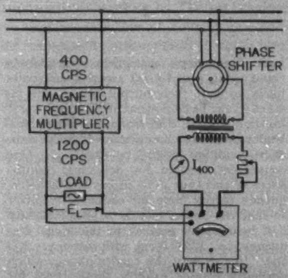


Fig. 5. Measurement of actual magnitude of the fundamental 400-cycle component of 1,200-cycle output voltage E_L of a magnetic frequency multiplier by means of a single-phase wattmeter, separately excited by a sinusoidal 400-cycle auxiliary current I_{400}

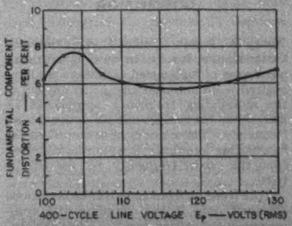


Fig. 6. Percentage fundamental component distortion of frequency-tripler output voltage E_L across the magnetic-amplifier circuit, as measured by means of the wattmeter method of Fig. 5

ponents did not exceed ±1/20 degrees, when introducing very large amounts of asymmetry into the push-pull circuit. For example, excessive variations of the bias resistors R_B' and R_B'' (about ±10 per cent) of the input-stage circuit did not have any effect on the angular position of the motor shaft. Thus, asymmetry zero-drift errors caused by very large deviations from symmetry of the push-pull circuit components can indeed be considered eliminated.

However, it is important to stress that the application of higher frequency ratio f_c/f_m, as described in this paper, does by no means decrease the well-known zero-drift effect of 400-cycle pickup voltages injected into the input circuit of the magnetic servo amplifier. Such additional error voltages, which may be caused by undesired capacitive currents or by unintentional coupling effects, should be eliminated by careful electrostatic shielding and grounding of windings and proper wiring design.[23]

Measurement of Bandwidth of Closed-Loop Servo System

Referring to the dynamic performance of the synchro-transmission system of Fig. 1, it is to be noted that the speed of response of the 2-stage half-wave, push-pull magnetic amplifier is inherently $1^1/_2$ cycles of its power supply frequency.[4,5] When supplying the magnetic servo amplifier with 1,200 cps, instead of 400 cps, actual speed of response will be 1.25 instead of 3.75 milliseconds. Results of various measurements on the closed-loop system using a servoscope arrangement have shown that the bandwidth of the system can be correspondingly increased in this way. For example, with the 2-stage magnetic amplifier of Fig. 1, measured bandwidth is about 19 to 23 cycles (120 to 145 radians), and up to 30 cycles (190 radians) have been measured with 3-stage half-wave circuits having an inherent speed of response of 1.67 milliseconds corresponding to 2 cycles of the 1,200-cycle carrier frequency.

Similar servoscope measurements concerning performance of synchro-transmission systems, supplied from the 60-cycle line and with the use of either a 60/180-cycle or a 60/540-cycle type of magnetic frequency multiplier, have shown that, when conventional 60-cycle 2-phase motors are used, bandwidth values up to about 10 to 11 cycles (63 to 69 radians) are obtained.

Evidently, besides the possibility of eliminating asymmetry zero-drift errors actual speed of response of the magnetic servo amplifier and the bandwidth of the closed-loop system will be appreciably increased.

References

1. Magnetic Amplifiers of the Balance Detector Type—Their Basic Principles, Characteristics, and Applications, W. A. Geyger. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1707–20.

2. Magnetic-Amplifier Circuits (book), W. A. Geyger. McGraw-Hill Book Company, New York, N. Y., 1954, chaps. 8, 9, 12–14.

3. Arrangement for Amplifying Electric Currents by Means of D-C Controlled Saturable Reactors, A. U. Lamm. *French Patent No. 863,397*, Jan. 2, 1941.

4. An Improved Magnetic Servo Amplifier, C. W. Lufcy, A. E. Schmid, P. W. Barnhart. *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 281–89.

5. Design Considerations of the Half-Wave Bridge Magnetic Amplifier, C. W, Lufcy, H. H. Woodson. *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 220–25.

6. Special Magnetic Alloys and Applications, G. W. Elmen, E. A. Gaugler. *Electrical Engineering*, vol. 67, Sept. 1948, pp. 843–45.

7. Method of Producing Grain Oriented Ferro-Magnetic Alloys, E. A. Gaugler. *United States Patent No. 2,569,468*, Oct. 2, 1951.

8. Magnetic Amplifiers with Orthonol Tape Cores, W. A. Geyger. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 6, 1950, pp. 59–67.

9. Frequency Multiplication—Principles and Practical Applications of Ferro-Magnetic Methods, N. Lindenblad, W. W. Brown. *AIEE Transactions*, vol. 44, 1925, pp. 491–96.

10. Frequenzwandler zur Erweiterung des Messbereichs von Frequenz-Messgeräten, W. A. Geyger. *Elektrotechnische Zeitschrift*, Braunschweig, Germany, vol. 44, no. 24, 1923, pp. 565–66.

11. Messung der höheren Harmonischen nach der Kompensationsmethode, W. A. Geyger. *Archiv für Elektrotechnik*, Berlin, Germany, vol. 18. 1927, pp. 634–38.

12. Magnetic Frequency Conversion, L. C. Harriott. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 9, 1953, pp. 78–87.

13. Magnetic-Amplifier Circuits (book), W. A. Geyger. McGraw-Hill Book Company, New York, N. Y., 1954, pp. 249–52.

14. Magnetic Frequency Multipliers, L. J. Johnson, S. E. Rauch, *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 448–52.

15. Frequency Multiplier, O. von Bronk. *United States Patent No. 1,678,965*, July 31, 1928.

16. Transformateurs statiques de fréquence. M. Joly. *La Lumière Électrique*, Paris, France, vol. 14, second series, no. 20, May 1911, pp. 195–204.

17. Static Frequency Transformer, M. Joly, *United States Patent No. 1,118,035*, Dec. 1, 1914.

18. Static Frequency Changer, F. Spinelli. *United States Patent No. 1,157,730*, Oct. 26, 1915.

19. Frequency Tripling Transformers, J. L. Cantwell. *AIEE Transactions (Electrical Engineering)*, vol. 55, July 1936, pp. 784–90.

20. Frequency Multiplier, H. M. Huge. *United States Patent No. 2,424,237*, July 22, 1947.

21. Laboratory Notes on Alternating-Current Circuits, W. E. Ayrton, J. Perry. *Journal*, Institution of Electrical Engineers, London, England, vol. 18, 1889, pp. 284–311 (p. 307: Analysis of a-c wave by means of electrodynamometer).

22. A Direct Method for Analyzing Alternating Currents, T. Des Coudres. *Elektrotechnische Zeitschrift*, Berlin, Germany, vol. 21, 1900, pp. 752–55; 770–72.

23. Trouble Shooting in Advance by Proper Wiring Design, G. Weiss. *Control Engineering* New York, N. Y., vol. 1, no. 1, Sept. 1954, pp. 49–54 (p. 49: "Trouble Effect of Pickup on a Synchro Servo").

Vol. 7
Pt. 1
1955

◆

Discussion

Henry W. Patton (Collins Radio Company, Cedar Rapids, Iowa): In working with various forms of high-speed amplifiers we have noticed a pronounced change in output as the modulating frequency approaches an integral or fractional multiple of the carrier frequency. These are most severe when the modulating frequency is one order different from the carrier frequency, i.e., one-half or twice.

True freedom from zero-drift errors can be obtained only if the "fundamental component in the carrier does not effect the output of the amplifier" as pointed out in the paper. If, however, this does occur, the effects could be minimized by locking the phase of the amplifier carrier to that of the motor supply and adjusting the phase relationships for minimum disturbance. Magnetic frequency multipliers, particularly those using tuned tank circuits, are quite likely to shift their output phase with respect to the basic power supply under varying temperature and frequency conditions. Under these circumstances an effective zero-drift condition might be expected. One convenient method for detecting disturbances of this type is to supply the high-frequency carrier from an oscillator-amplifier so that the carrier can be adjusted slightly "off beat" with respect to the signal frequency, i.e., 1,200.5 cps with respect to 400 cps. In this way the carrier phasor which normally rotates exactly three times for each rotation of the power supply phasor will gain phase slightly each revolution giving rise to a relative phase drift equivalent to 0.5 cps. The resultant interaction on the motor may be readily observed.

Such effects might cause nonlinearities each side of the null particularly considering interaction with the quadrature component found in synchro control transformers. The carrier system described will undoubtedly eliminate the major drift troubles with amplifiers, but we wonder if drift has been observed caused by phase shift as described.

Dr. Geyger is to be complimented on a worth-while contribution which gives us another tool with which to attack two difficult magnetic amplifier problems, time response and drift.

W. A. Geyger: To avoid the difficulties which would arise in the case that the modulating frequency (fm) is one order different from the carrier frequency (fc), i.e., one-half or twice, the magnetic-amplifier system described in the paper does employ a single-phase or 3-phase magnetic frequency tripler or a 2-stage type of magnetic frequency multiplier consisting of triplers connected in cascade for providing the suitable frequency ratio fc/fm = 3 or 9 respectively.

I am grateful to Mr. Patton for stressing that true freedom from asymmetry zero-drift errors can be obtained only if actual magnitude and phase relationship of the fundamental component in the carrier does not affect the output of the magnetic amplifier, as already pointed out in the paper.

Mr. Patton's suggestion of supplying the high-frequency carrier from an oscillator-amplifier so that the carrier can be adjusted slightly "off beat" with respect to the signal frequency offers a convenient method for detecting any disturbances which may be caused by an excessive fundamental-frequency output-voltage component of the magnetic frequency multiplier.

Field Tests on a 100-Megawatt Rectifier Installation

J. K. DILLARD
MEMBER AIEE

JOHN KIEFER
ASSOCIATE MEMBER AIEE

C. S. HAGUE
ASSOCIATE MEMBER AIEE

, 74
1
55

IN THIS paper results are presented from field tests made in March 1954 on the 100-megawatt ignitron rectifier station at the Robert P. Patterson plant, Arkadelphia, Ark., of the Reynolds Metals Company. Power from the rectifier station is used in the electrolytic reduction of aluminum, and all of the power is consumed on one aluminum potline. This is the largest concentration of d-c power on a single potline in this country.

During World War II, a number of rectifier installations were made to provide d-c power in the neighborhood of 60,000 amperes and 650 volts for production of aluminum and other light metals. As more efficient processes for making metal have been developed, there has been a trend toward higher voltages and currents on potlines built since the end of the war. Several potlines in this country now take 72,000 amperes at 750 volts. In January 1954, the 800-volt 125,000-ampere potline on which these tests were made was put in service at Arkadelphia.

The trend toward the use of larger rectifier installations has introduced new problems, principally with respect to switching larger blocks of power and providing for protection against higher arc-back currents. All of the phenomena associated with these problems have not been completely understood because of a lack of field data. These tests were made to obtain the data needed for a better understanding of these phenomena.

A 1-line diagram of the Arkadelphia rectifier station which consists of 24 ignitron rectifier assemblies, two assemblies being connected to each parallel double-wye rectifier transformer, is shown in Fig. 1. Test oscillograms were taken on two 9-element magnetic oscillographs mechanically coupled together, so that it was possible to record two oscillograms simultaneously. Klydonographs were located at strategic points to obtain records of surge overvoltages if any occurred. Both load-switching and fault-clearance tests were made. Since different principles are involved in these two switching operations, this report is divided in two parts: load-switching tests and arc-back tests.

Load-Switching Tests

Power interruptions on rectifier installations may occur either because of involuntary loss of a-c power or because of voluntary shutdowns, the latter often resulting from emergency conditions in the cell room. On older installations, where the maximum current is around 60,000 amperes, voluntary load drops are usually made by master-tripping the cathode circuit breakers. Where this practice is followed on large rectifier installations, load-dropping tests have shown that highly inductive high-current potlines impose severe interrupting duty on d-c circuit breakers.[1] The last breaker to clear must interrupt roughly one-half of the normal potline current[1] which also must be carried by the rectifier tanks and transformer for 1 or 2 cycles. The inherent limitations of dropping load by master-tripping systems were not overcome in those early electrochemical installations where individual a-c circuit breakers were used on each rectifier unit, because the cell line inductance forced high currents to shift to the last rectifier unit to clear, regardless of whether the operation was made by master-tripping the unit a-c breakers or by master-tripping the cathode breakers.

Some rectifier installations were originally set up to drop load by blocking the ignitor circuits to stop firing, but this method was abandoned because excessive overcurrents and arc backs frequently accompanied the switching operation.[1]

If load is dropped by interrupting the a-c supply to the high-voltage bus, and auxiliary power is maintained on the rectifier firing circuits, any anode is free to conduct and the decaying current to the potline may take any path through the d-c winding of the transformer. Current flow is confined to the low leakage inductance path, afforded by two anodes 180 degrees apart, and decays smoothly to zero in each rectifier unit. For this reason, this method of making voluntary trips, both planned and emergency, was recommended for large rectifier installations.[1] This is the method used at Arkadelphia. The purpose of the tests was to substantiate this practice.

TEST PROGRAM

The normal method employed at Arkadelphia for a planned trip of the potline is to run the voltage regulator down, apply roughly a 33-degree delay to the firing circuits, and then trip the three 13.8-kv line circuit breakers with auxiliary power kept on the rectifiers. For emergency trip, the same procedure is followed except the voltage regulator is not moved and no delay is applied. In restoring load following a power interruption, the sequence is to close the a-c line breakers, unshort-circuit the ignitor circuits, remove delay, and run the voltage regulator up to full position.

Tests were made to substantiate these practices for planned and emergency trips and for restoring load. In another test, for comparison, a 42-degree delay was used instead of the normal 33 degrees. Consideration also must be given to those situations where power interruptions result from involuntary loss of a-c power in which cases the source of auxiliary power is lost. Therefore, one additional test

---

Paper 55-39, recommended by the AIEE Electronic Power Converters Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31-February 4, 1955. Manuscript submitted October 21, 1954; made available for printing November 17, 1954.

J. K. DILLARD and C. S. HAGUE are with the Westinghouse Electric Corporation, East Pittsburgh, Pa., and JOHN KIEFER is with the Reynolds Metals Company, Arkadelphia, Ark.

The contributions of our colleagues who worked so hard in carrying out these tests are gratefully acknowledged.

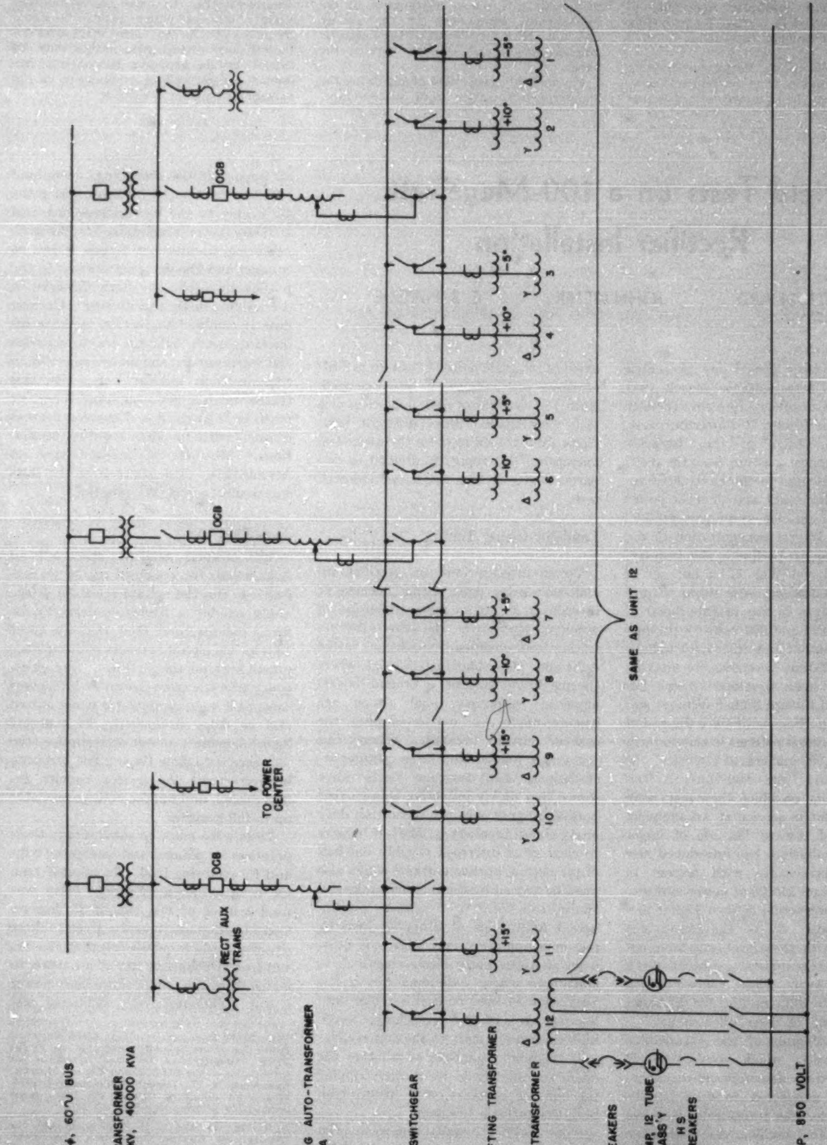

115 KV, 3 φ, 60 ∿ BUS
OCB
POWER TRANSFORMER
115–13.8 KV, 40000 KVA
REGULATING AUTO-TRANSFORMER
37,500 KVA
OUTDOOR SWITCHGEAR
PHASE SHIFTING TRANSFORMER
RECTIFIER TRANSFORMER
9,500 KVA
2–6 POLE
ANODE BREAKERS
2–5208 AMP, 12 TUBE
RECTIFIER ASS'Y
2–1 POLE HS
CATHODE BREAKERS
DC BUS
125,000 AMP, 850 VOLT
RECT AUX TRANS
TO POWER CENTER
OCB
OCB
OCB
SAME AS UNIT 12
1 Δ −5°
2 Y +10°
3 Y −5°
4 Δ +10°
5 Y +5°
6 Δ −10°
7 Δ +5°
8 Y −10°
9 Δ +15°
10 Y
11 Y +15°
12 Δ

Fig. 1. One-line diagram of Arkadelphia Plant

Table I. Initial Conditions—Load Switching Tests

Test No.	Volts	Potline Kiloamperes	Megawatts	Regulator Tap	Degrees Delay	Normally Used For
1*	640	78	50	1	33	Planned trips
2*	608	71	43	1	42	
3*	845	108	91	28	0	Emergency trips
4†	842	100	92	28	0	Loss of power
5‡		No load		1	15	Pickup load

*Trip a-c breakers with auxiliary power maintained.

†Trip a-c breakers and auxiliary power simultaneously.

‡Unshort-circuit ignitor circuits to pickup load.

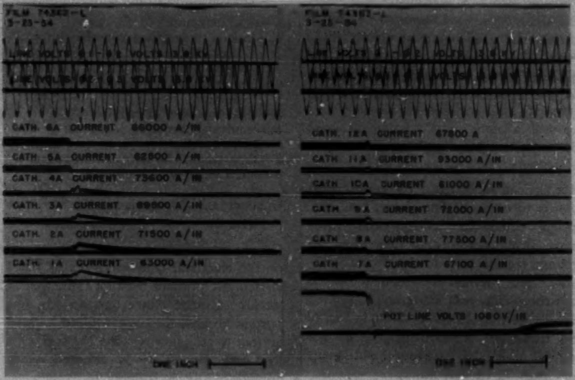


Fig. 2. Load-dropping test from full load with auxiliary power maintained, test 3

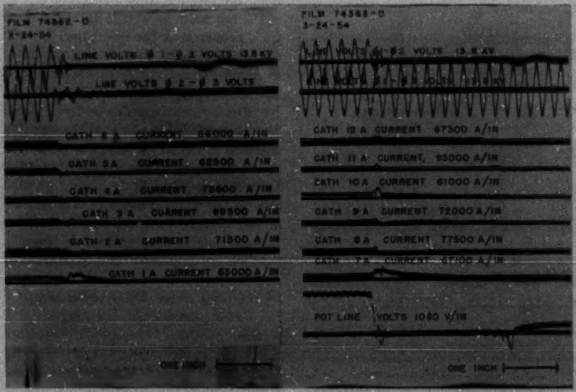


Fig. 3. Load-dropping test from full load with auxiliary power tripped simultaneously, test 4

was made by tripping the line circuit breakers and auxiliary power supply, simultaneously, with the voltage regulator in the full position and no delay was applied.

Initial conditions of the potline at the time of each test are tabulated in Table I. Readings were recorded from meters on the operator's control board just before the test. Oscillograms were made on each switching operation. There were not enough oscillographic elements available to record all 24 rectifier assembly cathode currents, but all of the *A* assembly units were recorded as well as potline voltage and voltages on the line side of the 13.8-kv circuit breakers.

RESULTS OF LOAD-DROPPING TESTS

Four load-dropping tests were made; the normal use of the method followed in each test is given in Table I. The oscillograms from test 3 are shown in Fig. 2. Oscillograms for the other tests where auxiliary power was kept on the rectifiers are not included here, because there were no essential differences in the results of the first three tests except in the magnitude of load dropped. The oscillograms from test 4 are shown in Fig. 3. To retain one a-c voltage trace for timing, the phase-1-to-phase-3 voltage was picked up back of the breaker used to drop the auxiliary power source. Maximum cathode currents are tabulated in Table II; these were measured from the oscillograms and are expressed in terms of their switching transient indexes (STI). STI is defined as the ratio of the maximum current after the switching operation to the current before the operation. For example, in test 1, each cathode circuit carried 3,250 amperes, since the load current of 78,000 amperes was divided among 24 cathodes. Thus, a STI of 2.0 in cathode 1*A* indicates that the maximum current transient was 6,500 amperes. This is just a convenient way of saying that this particular current doubled during the transient period.

The 13.8-kv bus at Arkadelphia is sectionalized, see Fig. 1, each of the three sections being supplied through an a-c circuit breaker. When load is to be interrupted, the three a-c breakers are tripped simultaneously, but all of them do not interrupt at exactly the same instant. Looking at the first three tests in Table II, where auxiliary power was maintained, the breaker supplying rectifiers 5, 6, 7, and 8 cleared first and these units immediately gave up their portion of the load to the other eight units. The breaker supplying rectifiers 9, 10, 11, and 12 was the second to clear, and the four units left on tried to hold the load. After a fraction of a cycle, the last circuit breaker cleared. The inductance of the potline prevents the sudden stoppage of current through it, and current trapped in rectifiers at the instant the a-c breakers cleared decayed exponentially to zero. In test 4, auxiliary power was tripped simultaneously with the a-c power supply. Since the auxiliary power source cleared ahead of the last line breaker, some erratic

Table II. Cathode Current-Switching Transient Indexes*

Cathode	Test 1	Test 2	Test 3	Test 4
1A	2.0	2.0	1.8	1.8
2A	2.0	1.8	1.8	1.0
3A	2.0	1.8	2.0	1.0
4A	2.0	2.0	2.2	1.0
5A	1.0	1.0	1.0	1.0
6A	1.0	1.0	1.0	1.0
7A	1.0	1.0	1.0	1.9
8A	1.0	1.0	1.0	1.5
9A	1.3	1.2	1.4	1.3
10A	1.4	1.4	1.3	2.0
11A	1.4	1.1	1.2	2.0
12A	1.1	1.4	1.2	1.3

*The ratio of maximum current following the switching operation to the current before the operation.

shifting of current among rectifiers was observed on this test.

A sectionalized bus arrangement is the most practical way to handle large blocks of power with equipment of reasonable size. The number of sections should be kept at a minimum consistent with practical equipment ratings, however, because of the tendency of current to increase on the last group of rectifiers to drop load. With three sections, as used at this plant, the transient current carried by any rectifier is not excessive.

Comparison of the first three tests of Table II shows that the introduction of delay before tripping the a-c breaker has no effect on the switching transient index. Delay is desirable in planned trips, however, to reduce the magnitude of load which is abruptly dropped from the power system.

Auxiliary power should be maintained when load is dropped on voluntary trips. The basis for this conclusion can be seen better by looking at Figs. 2 and 3 than at Table II; it is a matter of smoothness of current decay since no rectifier was overloaded beyond its short-time rating in either test. Therefore, in case of power failure where auxiliary power cannot be maintained, no dangerous overcurrents results.

Previous load-dropping tests[1] where master-tripped cathode breakers were employed showed switching transient indexes as high as 8.25, roughly four times the highest transient observed on any of these tests. This contrast vividly illustrates the advantage to be gained by making power interruptions on large rectifier installations by tripping the line circuit breakers.

Results of Load-Pickup Tests

The procedure at Arkadelphia for restoring load following a power interruption is to close the line circuit breakers, unshort-circuit the ignitor circuits, remove firing delay, and then run the voltage regulator up to the normal position. This method was used in test 5, and the oscillogram made as the ignitor circuits were unblocked is shown in Fig. 4. No transients occurred in any of the rectifier units, load current building up uniformly in each rectifier.

Nomenclature

E_b = counter electromotive force of the potline
E_d = terminal voltage of the power supply
E_d' = terminal voltage of power supply during a load trip
i = instantaneous potline current

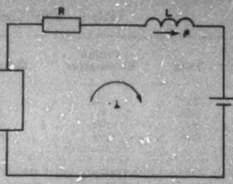


Fig. 5. Equivalent circuit representation of potline

L = total inductance of cell line
L_e = leakage inductance between d-c windings 180 degrees apart on rectifier transformer
n = number of double-wye assemblies in conduction
R = total resistance of cell line
R_a = equivalent arc-drop resistance
t = time, seconds
T = time, seconds for load current to stop
ρ = initial value of load current

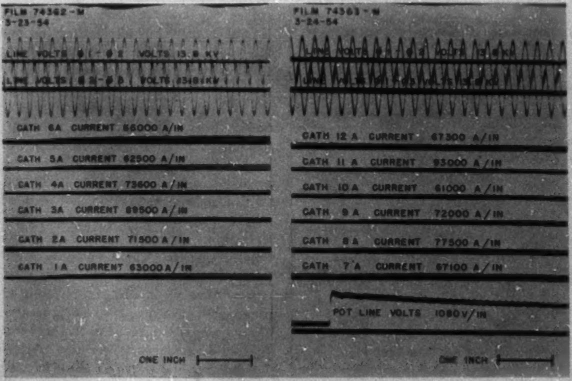


Fig. 4. Load pickup test, test 5

Electrical Characteristics of the Potline

Only limited data have been obtained on the electrical characteristics of fused electrolytic potlines used for reduction of metals. Circuit characteristics are important because the time required for load current to stop flowing after the start of a load-drop operation is affected quite appreciably by the inductance, resistance, and counter electromotive force of the cell line. These circuit constants depend upon the physical dimensions and layout of the cell room as well as upon the electrical properties of the metal, and it is not easy to determine them by direct measurement. Constants of the Arkadelphia line were obtained by relatively simple calculations based on the field test data. Similar data on other lines will have to be obtained if a clear picture is to be drawn of the range of constants for aluminum potlines in general.

A potline which is made up of a number of electrolytic cells may be represented by the equivalent circuit given in Fig. 5. Symbols are defined in the Nomenclature. The power supply, in this case 24 double-wye rectifier assemblies, is represented by the rectangular block. As long as any rectifier continues to pass current after the line circuit breakers open, the d-c busses are connected together by a circuit which includes the internal impedance of the rectifiers and transformers and the impedance of the d-c loop which connects each rectifier assembly to the d-c bus. When the line circuit breakers are opened, the terminal voltage of the power supply falls to a value determined by the impedance of this circuit. Since the largest impedances in the circuit are the arc-drop

Table III. Characteristics of the Potline

	p, Kilo-amperes	E_d, Volts	E_b/E_d	E_b, Volts	T, Seconds	R, Ohms	L, Milli-henrys
Test 1	78	640	0.334	214	0.150	0.0055	0.75
Test 2	71	608	0.345	211	0.142	0.0056	0.75
Test 3	108	845	0.250	211	0.192	0.0059	0.81
Test 4	109	842	0.250	211	0.192	0.0058	0.80
Average value						0.0057	0.77

resistance of the tanks and the leakage in-between the d-c windings of the transformer

$$E_d'=\frac{1}{2n}\left(R_a i+L_e\frac{di}{dt}\right)$$

approximately. If several assemblies continue to share load current, as they did on these tests, the parallel impedance of this circuit is very low and the d-c bus voltage E_d' drops practically to zero. When conduction stops, the bus voltage rises immediately to the polarization value, or back electromotive force, of the potline, as shown in Fig. 3.

Neglecting E_d' which is very close to zero, the solution of the circuit is given by

$$p=\frac{E_d-E_b}{R} \qquad (1)$$

and

$$i=\frac{E_d}{R}\epsilon^{-\frac{Rt}{L}}-\frac{E_b}{R} \qquad (2)$$

When $i=0$, i.e., when load current stops

$$\epsilon^{-\frac{Rt}{L}}=\frac{E_b}{E_d} \qquad (3)$$

Equations 1 and 3 can be solved for R and L since everything else in them can be measured from the test oscillograms. The numbers substituted in these equations and the resultant solutions are tabulated in Table III. For the Arkadelphia line: $R=0.0057$ ohm, $L=0.77$ millihenry, and $E_b=211$ volts.

Arc-Back Tests

An arc back results from the formation of a cathode spot on the anode of a rectifier tank which permits current to flow through the tank in the reverse direction. It provides simultaneous paths for short-circuit current from the secondary winding of the rectifier transformer and from the d-c bus. Consequently, arc-back currents are nearly always higher than those resulting from a 3-phase fault on the rectifier transformer. Knowledge of the magnitude of fault current is necessary for determining the interrupting duty on circuit breakers employed for arc-back protection.

Analytical calculation of arc-back current is a long and complicated procedure because several different modes of operation take place as other rectifier tanks on the faulted rectifier start and stop conduction.[2] For this reason, a study of arc-back phenomena was made on the analogue computer,[3] and empirical curves were developed which give the first crest of current in the anode and cathode circuits, providing the circuit constants are known. The first crest of fault current is of special interest because it occurs about the same time that high-speed circuit breakers start to limit the increase of arc-back current. It is a measure, therefore, of the maximum fault current which the equipment is subjected to, as well as the determinant of circuit-breaker interrupting duty.

The arc-back protection system at Arkadelphia consists of high-speed anode and high-speed cathode circuit breakers located as shown in Fig. 1. This is a little different from the conventional protection system employed on large rectifier installations in this country. Most large installations are protected by high-speed anode breakers, with current-limiting reactors on the a-c side of the rectifier transformer in those applications where the arc-back current can exceed the interrupting rating of available anode breakers

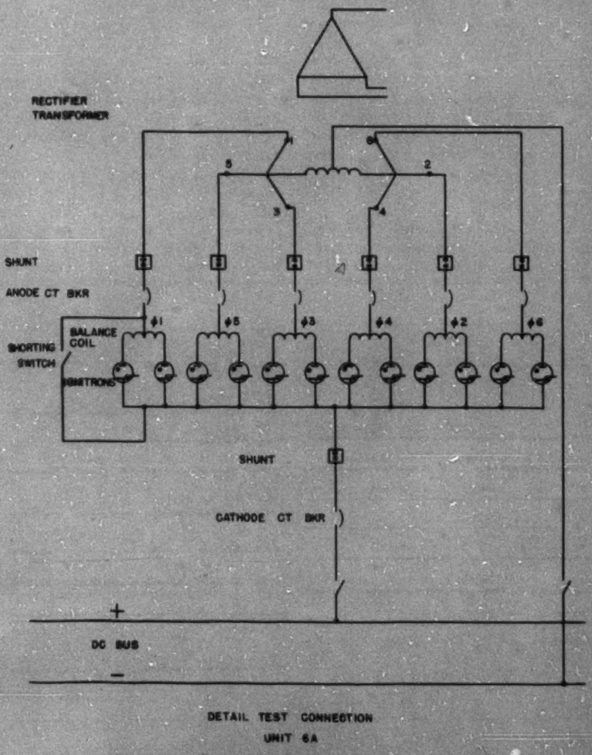


Fig. 6. Arc-back test circuit

Table IV. Initial Conditions—Arc-Back Tests

Test	Volts	Potline Kilo-amperes	Mega-watts	Time, Cycles*	Circuit
6	845	110	93	0.41	†
7	842	112	97	0	†
8	849	106	90	0.17	†
9	835	105	89	0	‡
10	835	105	90	0	‡
11	835	103	89	0	‡
12	835	102	90	0.40	‡
13	835	104	89	0	‡
14	835	105	90	0	‡
15	830	116	98	0.03	§
16	825	108	90	0	§
17	830	112	93	0.12	§

*Time in cycles from transition point to start of the arc back.

†Fig. 6 with 0.0018-ohm resistor in series with short-circuiting switch.

‡Fig. 6.

§Fig. 6 with d-c loop shortened.

or the fault current rating of transformers and other equipment. The conventional practice is to use semihigh-speed cathode breakers to isolate the rectifier assembly from the d-c bus. Use of high-speed circuit breakers in both the anode and cathode circuits is a departure from usual practice; at this plant, the magnitude of arc-back current permitted the combination to be used without reactors. Objectives of the tests were to compare fault currents obtained in the field with calculated values and to observe arc-back phenomena on a large rectifier installation.

TEST PROCEDURE

A short-circuiting switch was connected in parallel with the ignitron tanks phase 1 of rectifier *6A*, as shown in Fig. 6, and arc backs were initiated by closing the switch with the same cam used to operate the oscillograph. Oscillograms were made of all six anode currents on the test rectifier and cathode currents on the test rectifier *6A*, the adjacent rectifier *6B*, and the most remote rectifier *1A*. Potline current was also recorded, the current being measured by taking the voltage drop across a section of the d-c bus.

The arc-back protection system at the plant employs high-speed circuit breakers in both the anode and cathode circuits. These breakers work on a flux-shifting principle, and start to open whenever there is a rapid reversal of the flux set up by normal load current. Twelve tests were made in which arc backs were cleared by the plant protection system. Initial conditions of the potline, recorded from meters on the operator's control board just prior to each test, are tabulated in Table IV which also shows the circuit arrangement for each test.

In practice, most arc backs start at the end of a conduction period at the point of transition between the conducting and nonconducting intervals. Currents from arc backs which occur at transition are larger than those from arc backs which occur later in the cycle. Since complicated timing devices are required to hit exactly on a transition point every time in the field, the time at which the fault was applied was left to chance. This procedure is satisfactory because closing the short-circuiting switch anywhere in the conduction period is equivalent to starting the arc back at transition, as the voltage is not in the right direction to force reverse current to flow until the nonconducting period begins. The time, expressed in cycles of the fundamental 60-cycle-per-second frequency, between the transition point and the start of the arc back is given in Table IV.

Table V. Results of Arc-Back Test 11

	Test Value
Potline Conditions During Arc Back	
Voltage (per cent of initial)	78
Current (per cent of initial)	90
Reverse Current	
Phase 1 (crest kiloamperes)	64
Phase 1 maximum *di/dt* (amperes per microsecond)	11
Cathode *6A* (crest kiloamperes)	43
Cathode *6A* maximum *di/dt* (amperes per microsecond)	9
Forward Anode Current	
Phase 3 (crest kiloamperes)	21
Phase 5 (crest kiloamperes)	24
Phase 2 (crest kiloamperes)	0
Phase 4 (crest kiloamperes)	0
Phase 6 (crest kiloamperes)	16
Forward Cathode Currents	
Cathode *6B* (crest kiloamperes)	11
Cathode *1A* (crest kiloamperes)	7
Time in Degrees from Start of Arc Back to	
Crest current in phase 1	160
Crest current in phase 3	110
Crest current in phase 5	220
Crest current in cathode *6A*	160

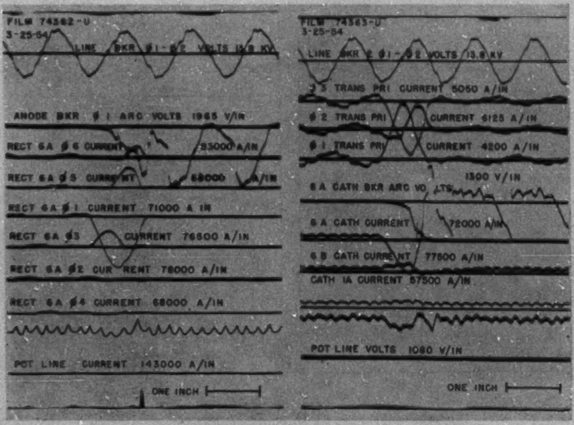


Fig. 7. Arc-back record, test 11

RESULTS OF ARC-BACK TESTS

In eight of the arc-back tests (tests 7, 9, 10, 11, 13, 14, 15, and 16), the fault was applied at, or very close to, the transition point. The oscillograms from test 11 are shown in Fig. 7, and values measured from them are tabulated in Table V. Oscillograms for the other arc-back tests are not included here because there were practically no differences in wave shapes or magnitudes between records on the eight transition shots.

The arc drop of the ignitron tanks under fault conditions limits the magnitude of arc-back current. In test 7, a 0.0018-ohm resistor was inserted in series with the short-circuiting switch to represent the equivalent arc-drop resistance of the faulted ignitron. The anode current in this test measured 62 kiloamperes, 3.1 per cent lower than in test 11 where the resistor was omitted. The reduction in magnitude appears about right, but it should be kept in mind that the observed reduction is comparable to the accuracy with which the oscillograms can be read. It was concluded that the resistor could be omitted as a conservative measure in subsequent tests without introducing significant error into the test results.

Another factor which influences the magnitude of fault current is the circuit consisting of the rectifier, its transformer, the interphase transformer, and the d-c loop (the feeders which connect the rectifier to the positive and negative d-c busses).

In tests 15 and 16, the area enclosed by the d-c loop was cut roughly in half by shortening the length of the connection from the interphase transformer to the negative d-c bus. Fault current magnitudes on these tests were practically the same as those tabulated in Table V, which substantiates the contention that most of the impedance to feedback current is presented by the inductance of the interphase transformer until it saturates.

Sympathetic arc backs are those which result from an arc back in another anode circuit. When they occur, the duration of the fault on the system is increased with corresponding increase in the disruptive effect of fault current stresses. The controlling factor in the production of sympathetic arc backs for any given ignitron design is the magnitude of forward-flow current fed into the original arc back by other ignitrons on the same wye group. In the tests, phase 5 carried crest currents of 24 kiloamperes divided between two parallel ignitron tanks. Not a single sympathetic arc back occurred on any of the 12 tests. Previous tests[4] have also shown that sympathetic arc backs are not likely to result from forward-flow currents of less than 13,000 amperes in ignitron tubes used in large electrochemical applications. The use of high-speed anode circuit breakers insures that forward-flow currents are kept at a minimum.

COMPARISON OF TEST AND CALCULATED VALUES

One objective of the tests was to compare magnitudes of arc-back currents obtained in the field with calculated values. The calculated values were obtained from empirical curves developed from an analogue computer study[3] and were based on the following circuit constants (see reference 3 for the curves and definitions of symbols):

$E_s = 750$ volts, line-to-neutral voltage on secondary of rectifier transformer
kva $= 4{,}750$, based on double-wye connection
$x_t = 7$ per cent, rectifier transformer reactance
$x_p = 1.5$ per cent, estimated reactance of regulating transformer plus high-voltage system
$L_d = 60$ microhenrys, inductance of d-c loop plus interphase transformer
$R_a = 0.0025$ ohm, equivalent arc drop of the ignitron during fault

Fault currents obtained from the curves are higher than those measured in the field for two reasons: the curves assume that the d-c bus voltage during the fault is 90 per cent of the prefault voltage, and limiting action of the circuit breakers is omitted in accordance with the standards for applying rectifier switchgear.

When an arc back occurs on a large rectifier installation there is a drop in d-c bus voltage caused by the voltage regulation of the parallel units feeding into the fault. The analogue computer curves[3] are based on the assumption that the bus voltage drops 10 per cent during an arc back. These tests show that a 27-per-cent drop in bus voltage is a more realistic figure at this installation. As the current in the cathode circuit is approximately proportional to the bus voltage, the value taken from the curves may be adjusted fairly accurately by multiplying by the factor (73/90). Since feedback current divides between the faulty anode and at least one normal anode on the same wye group,[3] the reduction of current in amperes through the faulty anode is about one-half the reduction in the cathode circuit.

Rectifier switchgear is applied on the basis of fault current which would be obtained if the circuit breakers were not present. In comparing field test magnitudes with calculated magnitudes of fault current, however, the limiting action of the circuit breakers used to clear the fault should be included in the calculations. In this case, the anode and cathode breakers were represented by resistances of 0.0050 ohm and 0.0025 ohm respectively. These values were obtained by taking the voltage across the breaker, corresponding to crest current, divided by the current. Such a representation assumes that the voltage developed by the circuit breaker is a linear function of the current, which is not strictly correct but close enough for comparison purposes.

Using the equivalent-resistance circuit-breaker representation with the circuit constants tabulated as in the foregoing, and following the procedure given in reference[3]

$$\frac{X_e}{R_e} = \frac{8.5}{2.6} = 3.3 \text{ and } \frac{X_s}{R_s} = \frac{8.8}{1.56} = 5.6$$

from which the following crest currents are obtained after adjusting the results from the curves for a 27-per-cent drop in bus voltage: anode circuit = 66,500 amperes; cathode circuit = 39,000 amperes. These check the measured values close enough for estimating purposes; see Table V.

KLYDONOGRAPH RECORDS

A great deal of study and exploratory testing has been devoted to the search for a source of transient overvoltages in rectifier installation.[1,4] Interest in this subject was generated by a few unexplained insulation failures at some earlier installations which appeared to have been caused by voltage stresses. However, no surge voltages have been indicated either on staged tests[1,4] or by the limited number of surge recorders which it has been possible to place in the field.[1] Nevertheless, the theory has persisted that transient overvoltages may be associated with switching operations, either for load dropping or clearance of faults. For this reason, during these tests, klydonographs were connected from each secondary terminal of the rectifier transformer to the neutral of the opposite wye, across the potline, and from phase to phase on the transformer secondary.

The highest voltage recorded in any location was 4,100 volts from the terminal of phase 5 to the neutral of the opposite wye group. Maximum voltages in the other locations ran from 2,000 to 3,000 volts. None of the voltages were high enough to damage the insulation provided on the rectifier transformer or other equipment. Similar results were obtained on previous tests.[1,4] With phenomena that are believed to occur seldom if at all, it is very difficult to draw definite conclusions from tests which give negative results. Nevertheless, the additional data from this investigation of transient overvoltages are reassuring.

Conclusions

Several conclusions may be drawn in the light of these field tests:

1. The best method for making voluntary load interruptions on large rectifier installations is by a-c line circuit-breaker operation with auxiliary power maintained on the rectifiers.
2. In case of power failure, high transient currents do not occur, but there is some shifting of currents between parallel rectifiers.
3. Where the a-c bus is sectionalized, the number of sections should be kept at the minimum consistent with equipment ratings.
4. No transient currents result from the procedure followed at this plant for picking up load.
5. The arc-back protection system consisting of high-speed circuit breakers in both the anode and cathode circuits showed excellent interrupting capability.
6. Calculated values of fault current obtained from analogue computer studies may be used with confidence.
7. With the ignitron design and protection system employed at this installation, the data indicate that sympathetic arc backs will not occur.
8. It is improbable that high surge voltages

are associated with circuit-breaker operations, either for load dropping or clearance of faults.

References

1. LOAD-DROPPING TESTS ON A LARGE IGNITRON RECTIFIER INSTALLATION, S. J. Pope, J. K. Dillard, C. R. Marcum. *AIEE Transactions*, vol. 72, pt. I, May 1953, pp. 164–75.

2. RECTIFIER FAULT CURRENTS—PART II, C. C. Herskind, A. Schmidt, Jr., C. E. Rettig. *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 243–52.

3. RECTIFIER ARC-BACK STUDY ON THE ANALOGUE COMPUTER, J. K. Dillard, C. J. Baldwin, Jr. *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 198–208.

4. ARC BACKS IN RECTIFIER CIRCUITS—ARTIFICIAL ARC-BACK TESTS, R. D. Evans, A. J. Maslin. *AIEE Transactions (Electrical Engineering)*, vol. 64, June 1945, pp. 303–11.

◆

Discussion

C. A. Langlois (Reynolds Metals Company, Richmond, Va.): The authors are to be commended for an excellent paper. Very little factual information has been available on the inherent characteristics of an aluminum potline.

The subject of load interruption to rectifier supplied aluminum reduction plants has been a subject of discussion for many years. However, the method of load rejection by opening the incoming a-c oil circuit breakers has been a more common practice than might be indicated from the paper even though factual data of the desirability of this method was lacking.

In 1941 the Reynolds Metals Company built a reduction plant at Longview, Wash., consisting of three 32,000-ampere lines. The manufacturer's original instructions for load interruption were to turn off the master excitation or to open the cathode breakers with the master trip circuit. The use of the master excitation method was ruled out for fear of damage to the equipment because of the extreme currents carried by the last tanks firing. The master cathode trip method was not always the cleanest interruption, for many times arc backs would result. There have been cases where units on other potlines which were not being interrupted have had arc backs simultaneously with the interruption.

Since in most cases the load was finally dropped by the opening of an a-c circuit breaker, a program of making required interruptions by various methods and observing the results was initiated. Following these tests instructions were issued in August 1942 that planned interruptions of the load were to be made by use of the incoming feeder oil circuit breaker with auxiliary power maintained. This presented some problems, and switching had to be done since both the rectifier auxiliaries and the plant distribution system are supplied through this breaker. In spite of this the interruption was to be made by opening the oil breaker, turning off the master excitation and then closing the oil breaker to restore plant power. The use of the cathode trips was still to be used for emergency tripping. In 1948 one of the cathode breakers flashed over into the control circuits and a complete outage of the plant resulted. Since that time the only method of potline interruption has been by the use of the oil circuit breaker. In 1951 additions to this station were undertaken to increase the load from 32,000 to 52,000 amperes per circuit. At this time the auxiliary power supply and normal mill power supply were removed from the potline incoming a-c breakers and master cathode tripping eliminated.

In 1946 the Reynolds Metals Company took over the operation of the Plancor plant at Troutdale, Ore. This plant has four 55,000-ampere potlines and differs from the Longview plant in the manner in which the general lighting and power load is supplied. In this case the mill load is supplied through a separate oil circuit breaker. However the auxiliary power to the rectifier station is supplied through the oil circuit breakers supplying the rectifier power. The master excitation method of load interruption was never used; however, the use of master cathode tripping is the accepted method for emergency tripping. All emergency load rejection, such as frequency relaying or potroom emergencies, trip the cathode breakers. This has never caused any material damage nor have arc backs been a difficulty. However, if time will permit, the load is to be rejected by first transferring the auxiliary power to another feeder and then opening the incoming oil circuit breaker. This method has been in use at the Troutdale plant since 1948.

Late in 1951, plans were made to construct the plant in Arkansas where the load-dropping tests of this article were made. Initially it was planned to provide for both a-c tripping and d-c tripping with preference to be determined during operation. Early in planning, on the recommendation of the manufacturer, the provision for d-c tripping was omitted.

No operating experience to date on lines being tripped from the a-c side indicates that this method is in any way less desirable than d-c tripping and, in addition to the advantages stated in the field test report, the gain in operating simplicity is quite important.

Lysle W. Morton (General Electric Company, Schenectady, N. Y.): It is gratifying to those who have strong interests in mercury-arc power rectifier equipment to take note of the rating of this rectifier installation of 100 megawatts. From the paper it is obvious that this large installation of mercury-arc rectifiers in the United States is a successful job. The mercury-arc rectifier deserves increased respect as a versatile and reliable type of a-c to d-c conversion apparatus. Much credit should go to the engineers of Reynolds Metals Company and Westinghouse for their over-all accomplishment in carrying out such a large power rectifier project.

The load switching tests and arc back tests reported in this paper proved to be informative and have increased the application know-how of those who apply large groups of mercury-arc rectifiers. The authors and their colleagues deserve hearty congratulations for what they have done.

In general, the conclusions drawn from the results of these field tests appear to be warranted. The shifting of current between parallel rectifiers in test 4, when the auxiliary power was dropped simultaneously with main a-c power, may give rise to questions as to the transient distribution of currents among the various anodes of a rectifier section and as to what may have occurred in the 12 rectifier circuits not being measured. This is no criticism of the test procedure. The problems of recording all phenomena in all circuits of such a complex circuit are well appreciated. We can't help wondering whether the observations and data taken during that one test are dependable indications of what will occur in every case of unplanned loss of a-c power.

The use of high-speed cathode breakers in combination with high-speed anode breakers was effective in controlling the crest arc back current and keeping it within standard limits of 75,000 amperes. This is one way to accomplish that objective, but other methods involving current-limiting reactance may prove as effective and less expensive.

In respect to the voltage surge question, we should like to know whether the rectifier transformers or rectifier circuits were equipped with surge suppression, and if so, with what type.

W. Fraser (Aluminum Laboratories, Ltd., Arvida, Que., Canada): This paper gives a very complete picture of conditions which take place during a drop or pick-up of load on large rectifier installations.

It appears, however, that the value given in Table III for potline inductance is far too low, and the error results from using the open-circuit voltage of the potline for the value of E_b. The value of E_b/E_d for a cell operating as its normal voltage is around 0.34 but this value for a series of cells varies with the number of cells which are on so-called anode effect. A potline with a number of cells on anode effect will reduce the E_b value of the line and increase the value of R.

When tests are carried out to determine the inductance of an aluminum potline, it is absolutely essential that all cells are in normal operating condition. The number of cells in this particular potline is not mentioned by the authors, but the value of E_b can be obtained fairly closely from tests 1 and 2 by extrapolation. Using this value for all four tests gives the following results in millihenrys (mh):

Test 1, $L=0.841$ mh
Test 2, $L=0.848$ mh
Test 3, $L=0.914$ mh
Test 4, $L=0.903$ mh

The average value of L would therefore be close to 0.88 mh. If 15 per cent of the cells were on anode effect during the higher voltage conditions, the recalculated values of L for tests 3 and 4 would be 0.885 and 0.844 mh respectively, the average value for the line being 0.855 mh. The values shown in Table III for current and voltage during tests 3 and 4 indicate abnormal conditions in the potline.

J. K. Dillard, John Kiefer, and C. S. Hague: Several interesting points have been brought out. The cases mentioned by Mr. Langlois are particularly enlightening because actual field experiences are cited relative to load rejection by several different methods: 1. removal of excitation; 2. master trip of cathode circuit breakers; and 3. a-c breaker operation. These experiences substantiate the conclusion given in this paper and in an earlier paper (reference 1 of the paper) that the preferred method of dropping load is by a-c circuit-breaker operation with excitation maintained. The a-c breaker referred to by Mr. Langlois and in the paper, pertains to the circuit breaker back of the common high-voltage bus which feeds several parallel rectifiers. This point is worth reemphasis at this time, because in some countries the practice is to use an individual a-c breaker on each rectifier unit. Where this is the case, the cell line inductance forces high currents to shift to the last rectifier unit to determine whether the operation is made by master tripping the unit a-c breakers, or by master tripping the cathode breakers.

When this test program was being planned, as Mr. Morton suggests, considerable thought was given to just what phenomena would be recorded in order to make the best use of the oscillograph elements available. Anode currents were omitted on the grounds that these had been recorded on similar tests conducted 2 years previously (reference 1 of the paper). The 12 cathode *B*-unit currents which were not measured were scattered uniformly along the line. Since results of the test at Arkadelphia as well as those at Mead (reference 1 of the paper), did not show any unexplained phenomena taking place in either the anode or cathode circuits, the results from test 4 are probably as dependable as can be obtained from a very limited number of tests. In answer to Mr. Morton's specific question, no surge suppression devices are employed on either the transformers or rectifier circuits at Arkadelphia.

Mr. Fraser's comments about the operation of aluminum lines on anode effect present a good understanding of some of the phenomena observed on tests on these large installations. However, none of the pots at Arkadelphia were on anode effect during any of the tests, all of the cells being in normal operating condition.

The value of E_b cannot be obtained by substituting the values given for ρ and E_4 for tests 1 and 2 from Table III in equation 1, and then solving the resulting equations simultaneously. This procedure requires the difference taken of two large numbers and, even though each may be accurate to a very few per cent, there may be very large errors in the difference. The back electromotive force of the Arkadelphia line is lower than that usually encountered. When the tests were made, there were 160 pots in service, the E_b of these particular pots being about 1.3 volts, more or less. Dividing 208 by the d-c bus voltage, 842 volts, gives $E_b/E_4=0.248$ which checks the value measured from the oscillograms and shown in Table V. It therefore appears that the value of R and L given in Table V are also fairly accurate and that the values of inductance suggested by Mr. Fraser are too high.

Values in Table III for current and voltage during test 3 and 4 were not taken from recording meters but were read from instruments located on the operator's board. The inaccuracy in reading these instruments accounts for the slight discrepancy between these two tests in the table. The potline was in normal operating condition during both tests.

A Communications System for the New York State Thruway

D. S. DEWIRE
ASSOCIATE MEMBER AIEE

H. A. EVANS
NONMEMBER AIEE

WITH present-day rapid transportation requirements for passengers and commodities, with increasing numbers of highway freight-carrying vehicles, and with vehicles for pleasure-driving within the economic reach of nearly everyone, the need for nonstop express superhighways has become most urgent. Many of these highways have been constructed and are in operation, and many more are in process of construction or in the planning stages.

One of the basic needs in operating such highways is instant communication between all operating forces to make the highway as safe as it is humanly possible, to expedite public vehicular movement, and, so that it may be self-sustaining, to facilitate collection of tolls and protection of the collectors. This, of course, involves communication between supervisory offices, collection booths, and all associated official vehicles on the highway.

Any discussion of the communications system provided for the New York State Thruway requires, first, a basic understanding of the administrative organization and operating procedures used and second, a description of the communications requirements and the technical means provided for meeting them. In the preparation of this discussion, therefore, close collaboration has been necessary between the Thruway Authority and the Upstate Area of the New York Telephone Company.

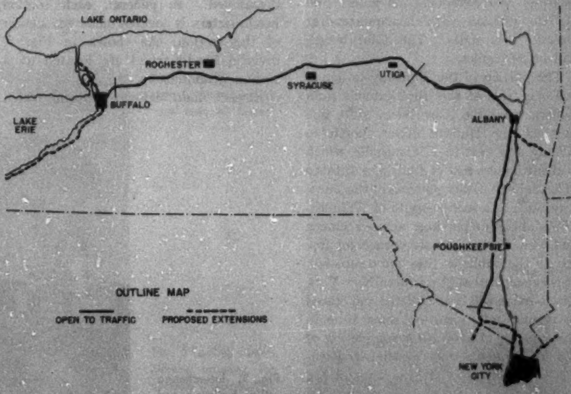


Fig. 1. Outline map of New York State Thruway

Paper 55-201, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing November 30, 1954.

D. S. Dewire is with the New York Telephone Company, Albany, N. Y., and H. A. Evans is with the New York State Thruway Authority, Albany, N. Y.

There have been many special problems encountered in designing this system. Meeting these problems and providing satisfactory communications for the Thruway has been accomplished through close understanding and co-operation between the New York State Thruway Authority, the New York State Police, and the New York Telephone Company, with invaluable assistance and advice from the American Telephone and Telegraph Company, the Bell Telephone Laboratories, and the Western Electric Company.

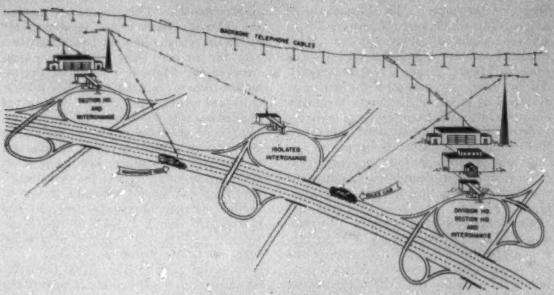

BACKBONE TELEPHONE CABLES
SECTION HQ AND INTERCHANGE
ISOLATED INTERCHANGE
POLICE CAR
DIVISION HQ SECTION HQ AND INTERCHANGE

Fig. 2. Typical elements of division network

Description of the Thruway

The New York State Thruway is a superhighway operated on a toll basis. As shown in Fig. 1, it will eventually extend from the southern Westchester County line (northern New York City Line) to the Pennsylvania state line near Erie, Pa. Its route lies north from New York City to Albany, west to Buffalo and thence southwest toward Erie. Branches are planned to Niagara Falls, and to the Massachusetts, Connecticut, and New Jersey state lines to connect with superhighways in those states.

It is planned ultimately to establish four administrative divisions of the Thruway as follows: New York division with headquarters at Elmsford on the east bank of the Hudson River near Tarrytown, 95 miles in length; Albany division with headquarters at Albany, 150 miles; Syracuse division with headquarters near Syracuse, 154 miles; and Buffalo division with headquarters at Buffalo, 163 miles. The total length will be 562 miles.

The portion of the Thruway in actual toll operation at this time extends from Suffern, near the New Jersey state line west of the Hudson River, north to Albany and west to Williamsville, which is about 5 miles east of Buffalo, a distance of 393 miles. Other portions of the Thruway, such as a section south of Williamsville and certain portions between Suffern and New York City, are opened for free use by the public as they are completed. It is expected that the Suffern-New York City portion will be completed and placed in regular toll operation some time in the summer of 1955 after completion of a 3-mile bridge across the Hudson River just south of Nyack in the Tappan Zee section.

To facilitate administration pending completion of the entire project, the Suffern-Williamsville portion has been divided into two administrative divisions with headquarters at Albany and Syracuse. The Albany division is 169 miles long and the Syracuse division is 224 miles long. The division boundary falls in the vicinity of Fort Plain on the southern bank of the Mohawk River. When the Suffern-New York City portion is ready for toll operation, it is expected that the New York division will be established, with the boundary between the New York and Albany divisions falling in the vicinity of New Paltz, across the Hudson River from Poughkeepsie. Eventually, a Buffalo division will be established with a new division boundary near Rochester.

Each division is subdivided into maintenance sections from 25 to 35 miles in length. Each maintenance section has a section headquarters from which highway maintenance crews and vehicles are dispatched. In general, each section headquarters is located near the center of the section. As shown in Fig. 2, division headquarters are located at a section headquarters.

At each main highway crossing, which has been selected as a point of access to the Thruway, a traffic interchange has been established. This is so arranged that traffic entering or leaving the Thruway passes through a "trumpet" arrangement in which the toll collection booths are located. Each section headquarters and division headquarters is located at one of these interchanges. The Syracuse division at present has 7 section headquarters and 14 isolated interchanges, or a total of 21 interchanges. In the Albany division there are 6 section headquarters and 9 isolated interchanges, making 15 interchanges in all. This totals 2 division headquarters, 13 section headquarters and 23 isolated interchanges (or 36 interchanges) now in operation.

Operation of the Thruway

In the operation of the Thruway, many types of official vehicles are used. Of major importance are police vehicles which patrol the Thruway for traffic control, accident assistance, reporting of hazardous road conditions, and general promotion of safety along the Thruway. The policing operations are handled by a special detail from the New York State Police organization.

Secondary to policing functions are:

1. Highway maintenance, with various supervisors' cars and highway maintenance equipment vehicles such as trucks, scrapers, and snow plows.
2. Highway automotive service, consisting of panel trucks and tow trucks. These trucks are furnished by commercial garages off the Thruway with which the Thruway Authority has contracted for this service. Each garage operator is responsible for 24-hour patroling of an assigned portion of the Thruway.
3. Ambulance and fire service, at present called from off the Thruway when required.
4. Toll collections, with toll collectors' cars making the run over the Thruway periodically to collect from the various interchanges.

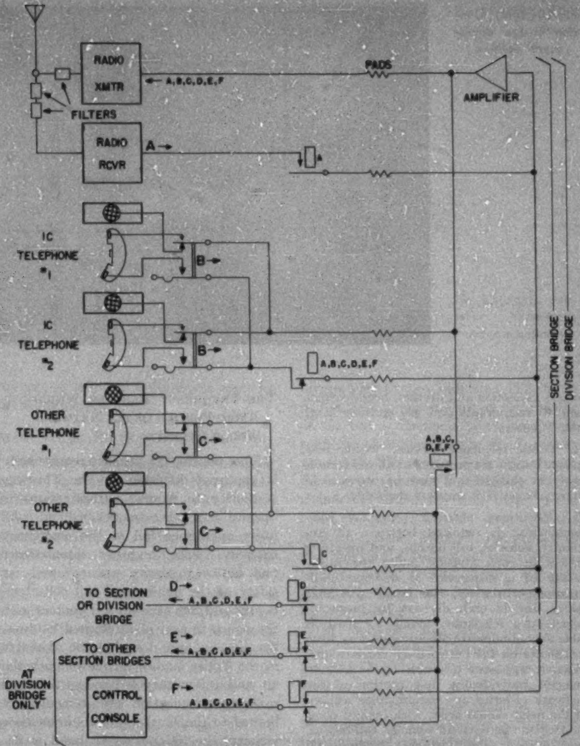
Fig. 3. Interchange utility building and toll booths

There are also various offices at the division and section headquarters and toll booths and a utility building at each interchange, which require fast and reliable communications service in the administration of the Thruway.

Communications Requirements

The communication requirements for the Thruway were outlined as follows:

1. Voice communication between authority headquarters and each division headquarters. At present, these voice communications requirements are met by the use of regular toll telephone service. It is anticipated that these requirements may grow to the point where it may prove economical for the Authority to arrange with the Telephone Company for the provision of a telephone channel for their exclusive use.
2. Teletype communication between Authority headquarters and division headquarters. At this time teletypewriter exchange or TWX service is in use. It is expected that this requirement may grow to the point where an exclusive-use teletypewriter channel may prove economical. This service is used primarily for administrative messages and for providing daily typed reports of collections in each division, but may be expanded to include weather reports, road conditions, and various other communications where a written record is desirable.
3. Telephone exchange service. Regular telephone service from the nearest telephone exchange is provided to all interchange toll booths and utility buildings, to


RADIO XMTR
← A,B,C,D,E,F
PADS
AMPLIFIER
FILTERS
RADIO RCVR
A →
A
IC TELEPHONE #1
B →
IC TELEPHONE #2
B →
A,B,C,D,E,F
OTHER TELEPHONE #1
C →
A,B,C, D,E,F
OTHER TELEPHONE #2
C →
C
SECTION BRIDGE
DIVISION BRIDGE
D →
TO SECTION OR DIVISION BRIDGE
← A,B,C,D,E,F
D
E →
TO OTHER SECTION BRIDGES
← A,B,C,D,E,F
E
AT DIVISION BRIDGE ONLY
CONTROL CONSOLE
F →
A,B,C,D,E,F
F

NOTE: PUSH-TO-TALK ORIGINATING AT "A", ETC., ACTUATES EQUIPMENT AT "A", ETC.

Fig. 4. Block diagram of division and section bridge equipment

Fig. 5 (left). Division headquarters and antenna tower

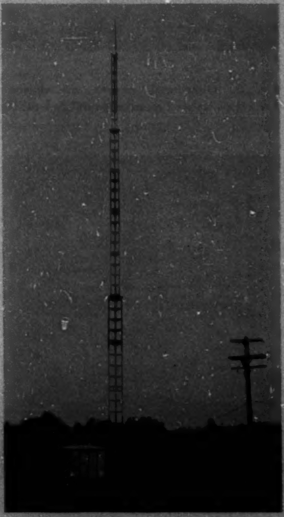


Fig. 6 (right). Temporary housing for section headquarters equipment, and antenna tower

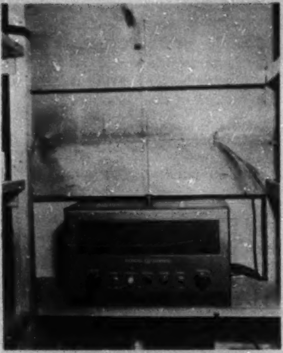


Fig. 7. Whip antenna inside isolated interchange utility building on radio equipment cabinet

Fig. 10 (right). Division bridge equipment cabinets

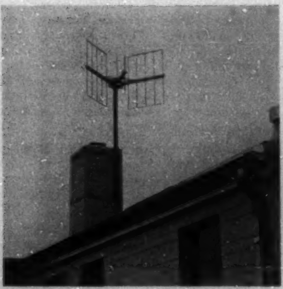


Fig. 8. Directional antenna on chimney bracket on isolated interchange utility building

Fig. 9. Section bridge equipment

offices in section and division headquarters, and to restaurants and gas stations along the Thruway.

4. Public telephone service. Public telephone booths are provided at all restaurants and gas stations and may be provided at interchanges if it appears desirable.

5. Operations network. Instant communication is required between all Authority vehicles, toll booths, and operating offices in each division under the supervision of a dispatcher at division headquarters. In effect, this means one huge party line in each division by means of which any telephone, mobile or fixed, is able to communicate directly with any other telephone on the network by voice calling. Also, it appeared desirable to be able to operate individually that portion of the network covering any maintenance section, if the need should arise, with instant communication between all vehicles, toll booths, and section offices in that section. Of course, in times of emergency, this communications channel is extremely busy but, by the use of rigid discipline as to message content and priority of communications under the control of a capable dispatcher, messages are so regulated that a minimum of delay results. Occasional delays are tolerable in view of the operating efficiencies of a single-channel system.

Communications Facilities

Requirements 1 through 4 are provided by the usual forms of common carrier service, and the facilities employed need not be discussed in detail here. The balance of this discussion will, therefore, be devoted to item 5, the operations network.

To connect vehicles to this operations network, radio was, of course, the only medium which could be used. The network, therefore, became primarily a radio system. In designing this network, many basic problems had to be solved. These problems and their solutions are as follows:

THE FREQUENCY BAND IN WHICH THE RADIO PORTION OF THE NETWORK WOULD OPERATE

Since policing is a primary requirement, it appeared desirable for the Thruway Authority to request authorizations to operate in a police frequency band. All other operations fall in the emergency category, since highway maintenance and highway service are required for safety of the public, and toll collectors' booths and cars need instant contact with the police in case of attempted holdups. Frequencies were found to be available in the Police Radio Service in both the 40- and 160-megacycle bands. Since all radio communication requirements are limited to the Thruway or its immediate vicinity, the 160-megacycle band is desirable to permit the use of economically feasible directional antennas where necessary. Further, this band is not as sus-

Fig. 11. Division bridge equipment

Fig. 12 (left). Dispatcher's control console

Fig. 13 (right). Isolated interchange bridge equipment

ceptible to skip and noise interference as the 40-megacycle band.

MAIN RADIO STATION LOCATIONS

For both operating and maintenance reasons, it was decided to locate main radio stations along the Thruway right-of-way. The distances separating proposed locations of section headquarters closely approximate calculated reliable coverage distances obtainable with reasonable radio transmitter powers and antenna support heights at these locations. The radio equipment is located in Thruway buildings, where heated quarters and emergency power are available. Access to the radio equipment is controlled by the Thruway Authority which is, of course, the licensee for the system. Since each section headquarters is located at an interchange, the Telephone Company maintenance people have access to the equipment location either from the connecting highway or from the Thruway.

It was necessary that these main radio stations be arranged as mobile relay stations (a 2-frequency system) so that communications between mobile units would be possible over an entire division. It was also necessary that they operate as base radio stations since they are controlled remotely from division headquarters and are directly connected to local telephones on the network in toll booths and section headquarters offices.

INTERCONNECTION OF MAIN RADIO STATIONS

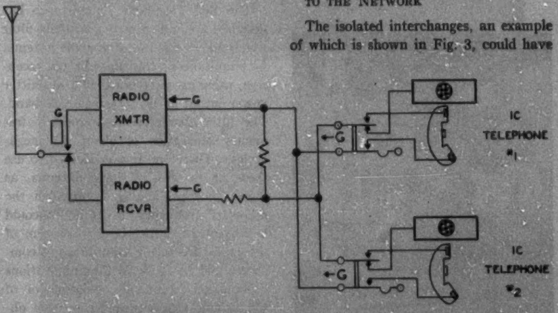
To provide one common network for an entire Thruway division, it is necessary that all main radio stations be interconnected for simultaneous operation. Although there were some advantages to be obtained by connecting these stations by microwave, studies indicated this to be neither as economical nor as dependable as the use of channels in existing main route or back-bone cables which closely parallel the Thruway. To be economical, microwave required a minimum of five channels in all parts of the system. Need could be foreseen for only three channels in some places. Further, the failure of one microwave station in the interconnection path would disrupt network communications until it could be placed back in service.

On the other hand, failure of any one back-bone cable is not serious since there are several cables along this route, paralleled by a main route microwave system. Further, in the unlikely event of failure of all back-bone routes, connection can be established readily from division headquarters to one or more section headquarters around the alternate main route between New York and the West via Pittsburgh and Cleveland.

In interconnecting the main radio stations, it is necessary to provide for individual operation of any section in an emergency without disturbing the rest of the division network. This was solved by carrying a channel from each main radio station back to a central point, at division headquarters. This permits provisions of an automatic selection circuit at division headquarters so that an incoming call via any section headquarters can lock out incoming calls from any other section headquarters. With microwave, this would have required as many as five channels in some sections, which means that a microwave equipment failure might isolate five sections at one time.

CONNECTION OF ISOLATED INTERCHANGES TO THE NETWORK

The isolated interchanges, an example of which is shown in Fig. 3, could have



NOTE: PUSH-TO-TALK ORIGINATING AT "G" ACTUATES EQUIPMENT AT "G".

Fig. 14. Block diagram of isolated interchange bridge equipment

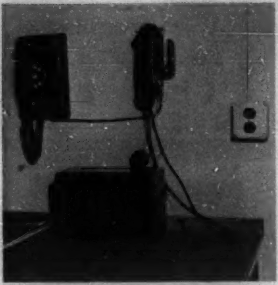


Fig. 15. Interchange utility building telephone equipment

been connected to the network by wire lines. This would have required the placing of many miles of new wire line facilities. It appeared economically desirable to locate a low-power base radio station at each such interchange to operate into the network in the same manner as mobile stations. This has proved a most satisfactory solution.

Connection of Section Headquarters Telephones to the Network

Since the toll booths, utility building, and section offices local to each section headquarters are to be in buildings reasonably close together, it is relatively simple to interconnect all network telephones in these buildings by cable with a specially designed bridge equipment, as shown in Fig. 4. The main radio station at that point is made a part of the bridge.

Connection of Division Headquarters Telephones to the Network

This is accomplished in the same manner as the section headquarters, with the addition of a control console for the division network dispatcher.

Fig. 16. Interchange booth

Radio System Transmission Design

In establishing main radio stations, theoretical computations were made based on studies of topographical maps. These computations were tempered by the wide experience of the Telephone Company engineers in mobile radio performance throughout the general area of the Thruway. In questionable cases, field tests were made to determine the validity of the computation-experience results and modifications were made in accordance therewith. The results of the computations and tests determined the required antenna height at each main radio station location, the radiated power which would be required, and the type of antenna to be used. Antenna heights vary from 75 to 135 feet above ground level at the radio station sites.

Communication between vehicles required use of a 2-frequency system with each main radio station designed to receive simultaneously a call on one frequency and retransmit it on another. There were several ways in which this could be done. Two antennas horizontally spaced by at least 200 feet would have been satisfactory. This would have meant two antenna supports from 60 to 120 feet in height. The supports had to be free-standing, since the design of roadways and buildings at each section headquarters precluded the use of guys. Two structures were undesirable both from space and economic standpoints. A single antenna support with antennas vertically spaced by at least 20 feet would have been possible, but this would have increased the height of the supports from 80 to 140 feet, which was also undesirable from an economic standpoint.

The Bell Telephone Laboratories has designed a small, compact antenna filter which allows the use of a single antenna for simultaneous transmission and reception, using up to 60 watts of transmitter power output in the 160-megacycle band. The filter characteristics require a frequency separation of about 4.5 megacycles. Therefore, it was decided to use these filters with a single antenna, as shown in Figs. 5 and 6, provided the necessary frequency assignments could be obtained. With the co-operation of the police frequency co-ordinating committee and the Federal Communications Commission, frequency assignments of 154.65 and 159.15 megacycles were obtained by the Thruway Authority.

Because of the power-handling capabilities of the new filters and previous experience, it was desirable to limit the power output from any main radio station transmitter to 60 watts. With the antenna heights mentioned in the foregoing and the use of gain antennas, this has been generally adequate. However, at each of two locations, subsequent operating experience has dictated the use of a 250-watt transmitter, with one cavity filter having greater power-handling capabilities.

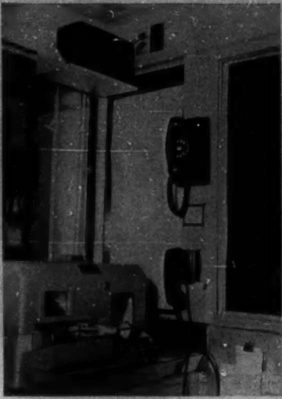


Fig. 17. Interchange booth telephone equipment

It was originally planned to direct antenna radiations along the Thruway. This would have required the use of two directional antennas, with maximum radiation in opposite directions, at each main radio station. An equivalent or better gain in antenna radiation could be obtained by the use of a single omnidirectional gain antenna, although radiations would then not be confined to the Thruway. The latter has proved generally desirable, since it is necessary to reach State Police barracks off the Thruway where the police who patrol the Thruway are quartered. If in the future

Fig. 18. Police car

Fig. 19. Police car radio equipment

Fig. 20. Police car radio control equipment

it becomes necessary, because of possible interference with other police systems, to direct radiations along the Thruway, the omnidirectional antennas can be replaced with directional arrays.

There are two or three cases where a section headquarters is near one end of a section. Here a single corner reflector is used rather than a gain antenna to cover that section. There is at least one case where two corner reflectors are used rather than a gain antenna to minimize the possibility of interference from other police systems using the same frequencies.

Calculations indicated that a low-power radio station would be sufficient at any isolated interchange. At isolated interchanges within a few miles of a base and mobile relay station, a whip-type antenna on the radio cabinet, as shown in Fig. 7, has proved adequate. For isolated interchanges at a greater distance, a corner reflector antenna is used mounted on a chimney bracket on the utility building, as shown in Fig. 8, directed toward the nearest base and mobile relay station. This makes it possible to use the low-powered radio equipment and also prevents hash in incoming messages where the signal strength from two adjacent base and mobile relay stations might be nearly equal. These isolated interchange radio stations have been classed by the Federal Communications Commission as base stations while the main radio stations are classed as base and mobile relay stations.

The radio transmission objective was 100-per-cent satisfactory coverage along the route of the Thruway. With the transmitter powers, tower heights, and antennas provided this objective has been realized. Hash areas exist where transmissions to a mobile unit are distorted because of phase differences from two adjacent transmitters in equal signal strength areas. However, with the system design employed, these areas are very short and, in general, have little effect on intelligibility. Occasional cases of more severe distortion present no particular problem, since on a report of unintelligibility the dispatcher can disable one of the two transmitters and repeat the message over the other. Because of the selective features of the division bridge, hash on calls incoming from mobile units or isolated interchanges does not exist, since the selected receiver locks out all others.

BRIDGE EQUIPMENT DESIGN

It was mentioned that the telephones and radio equipment at each section headquarters are interconnected by means of a bridge. The requirements to be met by this bridge equipment were:

1. Reradiation, on the main radio station transmitter frequency, of incoming calls from mobile units or isolated interchanges.
2. Reception of all calls at all telephones on the division network.
3. Voice calling.
4. Prevention of interruption to transmissions in progress.
5. Assignment of a control station, in accordance with the Federal Communications Commission rules, which would have a "carrier on" indication to show when the radio transmitter control circuits were energized, and a radio transmitter cutoff key to disable the radio transmitter in case of trouble which would produce unauthorized radiations.
6. A radio receiver cutoff key, readily available in case of trouble in the receiver of presence of an external noise condition

Fig. 21. Thruway passenger car radio equipment

Fig. 22. Dump truck

Fig. 24. Dump truck radio control equipment

which might lock the receiver in and cause noise over the entire network.

7. A push-to-talk handset and a loudspeaker with individual amplifier at each telephone.

8. At local interchange telephones, the ability to talk and receive on the network and also the ability to intercommunicate between themselves without affecting the network.

9. At all other telephones the ability to communicate on the network only.

The bridge equipment at each section headquarters is illustrated in Fig. 9. The heart of this bridge circuit is, as shown in Fig. 4, an amplifier which accommodates an outgoing circuit to the radio transmitter, an incoming circuit from the radio receiver, a 4-wire line circuit on which all interchange telephones are bridged, a 4-wire line circuit on which all other telephones are bridged, and a 2-wire line circuit to the division bridge. The amplifier is provided with an input bus and an output bus both designed for low impedance to minimize effects of trouble on any connected circuit. To provide the push-to-talk control feature, each radio receiver is equipped with a carrier-operated relay, and a control relay was provided in each telephone circuit. Push-to-talk control is also transmitted by means of control relays and a telegraph channel from the section bridge to the bridge at division headquarters. On a transmission in progress, all push-to-talk control is locked out except from the originating telephone and the dispatcher.

The bridge equipment at division headquarters is illustrated in Figs. 10 and 11. This bridge has the same requirements as the bridge at each section headquarters with certain additions and exceptions. In addition to the other telephones at section headquarters, there are other telephones in the division offices, and a network dispatcher. The dispatcher is required to have certain controls for the entire network. He can interrupt and take over the network from any other fixed or mobile telephone in case of emergency by merely depressing his push-to-talk button. He can disable the local transmitter and the local receiver. He has a "carrier on" indication for his local radio transmitter and also for all other main radio transmitters in the network as long as they are operating on a complete network basis. He can disable the radio transmitter at each remote main radio station as long as this station is operating on a network basis. He can cut off each line to a remote main radio station so that, in case of line trouble, the station affected can operate independently from its local control telephone. The dispatcher operates from a microphone. All other telephones, including those in mobile units, are equipped with handsets.

Fig. 23. Dump truck radio equipment

As with the section bridge, the heart of the division bridge is a main amplifier. Circuits are terminated on this amplifier in a manner similar to those at a section bridge, with the addition of a line circuit for the dispatcher and line circuits for all remote section bridges. A small console, shown in Fig. 12, is provided for the dispatcher, to accommodate a loudspeaker and amplifier, microphone preamplifier, control keys and associated guard lamps, and the "carrier on" lamp.

A small, relatively simple bridge is provided at each isolated interchange, as shown in Figs. 13 and 14. The requirements at this bridge differ considerably from those at the section and division bridges. Neither reradiation nor prevention of interruption, or lockout, are required. It has been necessary to assign a particular telephone as a control station, and to provide at this telephone a "carrier on" lamp and transmitter cutoff key. A receiver cutoff key is located in the utility building. All telephones require access to the network and also intercommunication without disturbing the network. Since each loudspeaker at a telephone has its own individual amplifier intercommunication is possible without

Vol. 7
Pt. 1
1955

additional amplification. Interchange telephone equipment is illustrated in Figs. 15, 16, and 17.

Control features at the various bridges are of interest. At a section bridge, the local radio receiver is equipped with a carrier-operated or codan relay, which operates whenever an incoming call is received from a mobile unit or nearby isolated interchange and provides a push-to-talk control signal. This signal disables the push-to-talk control from all local wire-connected telephones, activates the associated radio transmitter, and places a control signal over the control channel from the section bridge to the division bridge. At the division bridge, this control signal disables the push-to-talk control from all wire-connected telephones except the dispatcher, disables the codan control from the division radio receiver, activates the division radio transmitter, and is transmitted out over the control channels to all other section bridges. At the other section bridges, receipt of this control signal disables the push-to-talk control from all local wire-connected telephones, activates the section radio transmitter and disables the control from the section radio receiver. Disabling of all push-to-talk controls except that of the originating telephone prevents interruption by any other party except the dispatcher. Likewise, with a clear channel, depressing the push-to-talk button at any wire-connected telephone at a section bridge disables all other controls, activates the local radio transmitter, and passes the control signal to the division bridge and all other section bridges.

Operation at the division bridge is similar except that the dispatcher's control is never disabled. The dispatcher may, at any time, by depressing his push-to-talk button, disable all other controls and activate all main radio transmitters. This is accomplished by use of a single control channel between each section bridge and the division bridge which is a reversible channel to transmit control in either direction. In addition, this same control channel, by the use of polarized relays and reversed battery, is used for remote transmitter cutoff under control of the dispatcher.

Mobile Units

The mobile units used in this system, illustrated in Figs. 18 through 24, transmit on the main radio station receiver frequency and receive on the main transmitter frequency. The transmitters are of 25-watt power output, which has proved adequate. The noise conditions at the cars are somewhat more severe than those at the main radio station, so that 25 watts at the cars and 60 or 250 watts at the main stations provide a well-balanced system. Mounting arrangements in the cars were somewhat of a problem but have been worked out satisfactorily for the various types of vehicles used. The police cars, illustrated in Fig. 18, had a special problem since it was required that these cars have radio units on both the Thruway channel and the State Police channel. The State Police channel utilizes frequencies in the 40-megacycle band. Therefore, it was necessary to install two complete units. The transmitter and receiver chassis are placed on the shelf in the luggage compartment, as shown in Fig. 19. Both control units are located under the dash in the front, Fig. 20, with both loudspeakers under the grilled opening in the dash. The 40-megacycle antenna is mounted on the left rear fender, and the 160-megacycle antenna in the middle of the roof, as shown in Fig. 18.

Special Feature

One interesting special requirement concerned emergency intercommunication between the Thruway dispatcher and State Police Troop Headquarters along the Thruway. A monitoring receiver on the State Police channel is installed adjacent to the dispatcher's console, as shown in Fig. 12. Because of the nature of the wire network connecting the radio stations of the State Police system, a Troop Headquarters can activate the State Police radio station nearest the Thruway division headquarters and be heard on the monitoring receiver. Likewise, a monitoring receiver on the Thruway channel is installed in the Troop Headquarters, and all communications over the Thruway channel are heard on it.

Several other special features have been found desirable by the Thruway Authority since this system has been placed in operation. Provision of these features has presented new design problems which are being worked out for application to the system. When complete, it is expected that the Authority will have one of the most comprehensive and efficient communications systems ever provided.

Coaxial Cavity Filters for Multiplexing of 900-Megacycle Radio Relay Systems

M. H. KEBBY
ASSOCIATE MEMBER AIEE

Synopsis: Band-pass filters have been developed to provide for simultaneous application to a single antenna of two or more radio transmitters and/or receivers operating in the 900-megacycle (mc) range. By multiplexing transmitters or multiplexing receivers, or by diplexing a transmitter and receiver, considerable saving can be realized in equipment cost for multichannel radio systems. Filters consist of one or more sections of coaxial transmission line in tandem, adjusted for the desired degree of selectivity. In addition, filters are used to provide preselection and harmonic suppression where separate antennas are used for each transmitter and receiver.

ANTENNAS used with microwave radio systems are normally capable of handling efficiently a much wider band of frequencies than is required for a single, unidirectional group of communication channels. In this respect, antennas can be compared to open-wire lines or cable pairs, both of which are capable of transmitting a wider frequency band than is occupied by a single telephone channel or group of channels in a particular part of the frequency spectrum. Just as physical lines are commonly multiplexed to use fully their transmission band capabilities, so are microwave antennas used for simultaneous transmission and/or reception of more than one group of radio channels. Physical lines are multiplexed by using coil-capacitor or, in some cases, crystal filters to separate the various frequency bands involved. Although such conventional filters are not practical at micro-

---

Paper 55-200, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing December 7, 1954.

M. H. Kebby is with the Lenkurt Electric Company, Inc., San Carlos, Calif.

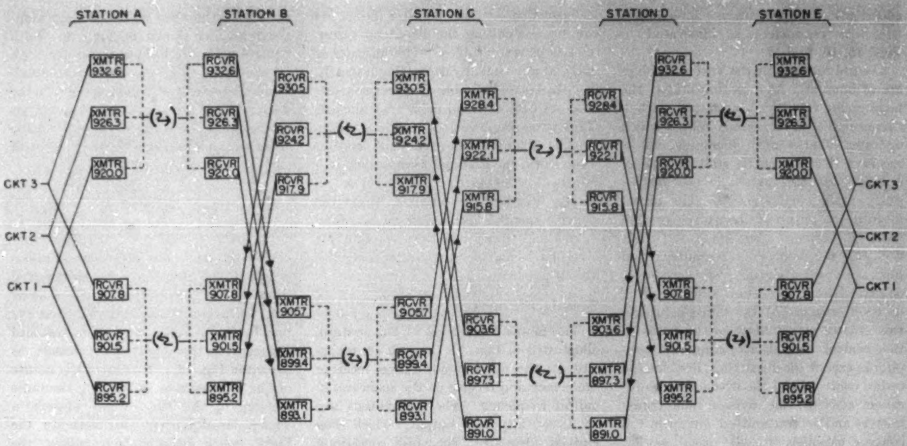

STATION A
STATION B
STATION C
STATION D
STATION E
XMTR 932.6
XMTR 926.3
XMTR 920.0
RCVR 907.8
RCVR 901.5
RCVR 895.2
RCVR 932.6
RCVR 926.3
RCVR 920.0
XMTR 907.8
XMTR 901.5
XMTR 895.2
RCVR 930.5
RCVR 924.2
RCVR 917.9
XMTR 905.7
XMTR 899.4
XMTR 893.1
XMTR 930.5
XMTR 924.2
XMTR 917.9
RCVR 905.7
RCVR 899.4
RCVR 893.1
XMTR 928.4
XMTR 922.1
XMTR 915.8
RCVR 903.6
RCVR 897.3
RCVR 891.0
RCVR 928.4
RCVR 922.1
RCVR 915.8
XMTR 903.6
XMTR 897.3
XMTR 891.0
RCVR 932.6
RCVR 926.3
RCVR 920.0
XMTR 907.8
XMTR 901.5
XMTR 895.2
XMTR 932.6
XMTR 926.3
XMTR 920.0
RCVR 907.8
RCVR 901.5
RCVR 895.2
CKT 3
CKT 2
CKT 1
CKT 3
CKT 2
CKT 1

Fig. 1. Frequency allocation for operation of three type-72B-900 radio systems over one path

Frequencies shown with transmitter and receiver assemblies in mc

wave frequencies, other types of frequency selective devices can be used.

In the 900-mc region, coaxial cavities are the most practical and economical type of filters, largely because of their excellent electrical characteristics and convenient size. Such accessory equipment is used in the recently developed Lenkurt type-*72B* radio systems. Although designs available at present use short-circuited quarter-wave coaxial sections tuned for operation only in the 900-mc band, similar filters are practical for other frequency ranges.

In possible applications of type-*72B* radio equipment, it is anticipated that up to three parallel radio circuits may be required in the 900-mc band (890 to 936 mc), and a frequency allocation permitting such operation has been established, as shown in Fig. 1. After allowances for such factors as cross-coupling, overshoot, and spurious responses, filters have been designed with sufficient selectivity to permit a minimum separation of 12 mc between any transmitter and receiver operating on the same antenna. Minimum separation between two transmitters or between two receivers on a single antenna is 6.3 mc.

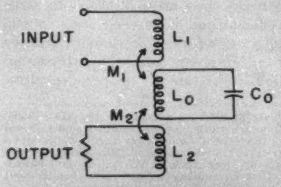

INPUT
L1
M1
L0
C0
M2
OUTPUT
L2

Fig. 2. Approximate equivalent circuit of coaxial cavity filter near resonance

Coaxial cavity filters acting as multiplexers, diplexers and multidiplexers have been developed for the following equipment combinations.

1. One transmitter and one receiver: diplexer.
2. Two transmitters:
3. Two receivers:
4. Three transmitters:
5. Three receivers:

} multiplexers

6. Two transmitters and two receivers: multidiplexer.

Coaxial Cavities as Radio-Frequency Filters

Coaxial cavity filters consist of one or more short segments of coaxial transmission line with each segment corresponding to an individual mesh of a conventional coil-capacitor filter. Each line section is effectively a quarter wave long at the center of the passband (referred to as the resonant frequency); it is short-circuited at one end and open-circuited at the other, with coupling loops provided for the input and output connections. In the vicinity of resonance, the impedance-frequency characteristics of such a line are equivalent to those of a lumped-constant parallel-resonant circuit with an equivalent Q. Generally, the desired attenuation characteristics of the radio-frequency filter lie in the frequency range of about 5 per cent above and below the resonant frequency. Therefore, within this range, standard response curves relating bandwidth and Q derived for lumped-constant circuits may be used.

The frequency of resonance of a transmission line in mc is given approximately by $f=246n/L$, where n is the number of quarter-wave lengths along the line and L is the line length in feet. In a radio-frequency filter, a line is ordinarily used at its lowest possible (fundamental) resonant frequency, in order to have the greatest possible frequency separation between higher-order resonant modes. With a short-circuited quarter-wave line, the closest resonant frequencies occur at odd multiples (3, 5, 7, etc.) of the fundamental frequency, while with a half-wave open line the higher-order modes

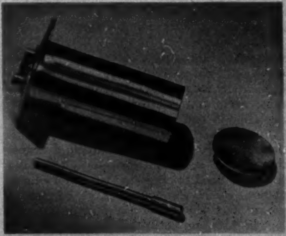


Fig. 3. 900-mc coaxial cavity filter

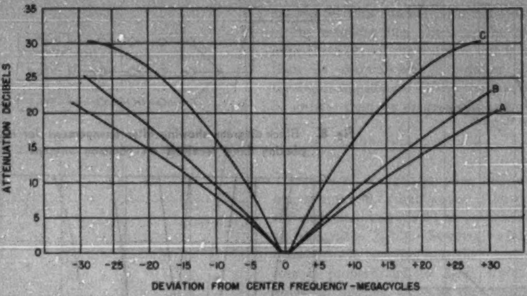


Fig. 4. Attenuation versus frequency characteristics of single-section coaxial cavity filter for several coupling conditions

are even multiples of the fundamental. In addition, the quarter-wave line is shorter and occupies less space, an important consideration in the ultrahigh frequency and lower microwave spectrum. For these reasons, a quarter-wave line was selected for this design.

Coupling Methods

Various arrangements can be used for coupling energy into and out of tuned cavities. The coupling may be either magnetic or electric. Coupling to the magnetic field is most effective near the short-circuited end of a quarter-wave line where the current density is greatest, while coupling to the electric field is most effective near the open-circuited end where the highest voltage gradient exists. Magnetic coupling may be obtained by placing a small loop of wire or ribbon (which is connected to a coaxial transmission line) in the cavity. Electric coupling is usually accomplished with a small capacitive probe, formed by extending the center conductor of the coaxial transmission line into the cavity.

Magnetic loop coupling has been chosen as the most practical method in the present design. The loops protrude through the short-circuiting plate of the coaxial cavity. Input and output loops are placed on opposite sides of the center conductor to provide minimum direct magnetic coupling between the loops. Such coupling between loops would tend to reduce the selectivity of the filter.

The effect of coupling loop position may be seen by considering an approximate equivalent circuit of a cavity filter near resonance (Fig. 2). The loop inductances (L_1 and L_2) will be tuned out by the resonant circuit (L_0 and C_0); however, they should be kept as small as possible to prevent altering of the impedance characteristics of the resonant circuit.

The mutual coupling (M_1 and M_2) between the loops and the resonant circuit is dependent upon the loop areas and their position and orientation in the magnetic field. The input impedance at resonance is dependent primarily upon the relationship between M_1 and M_2 (for a given load impedance). Generally, M_1 is made equal to M_2, so that at resonance, the input impedance is approximately equal to the load impedance. At other frequencies, the input impedance is dependent more upon the impedance of L_0 and C_0 which, of course, drops as the frequency deviates from resonance. This in turn results in an extremely low input impedance which is practically independent of the load impedance.

The rate at which the input impedance drops as the frequency deviates from resonance (selectivity of the filter) is dependent upon the loaded Q of the resonant circuit; this is determined primarily by the value of M_2 (for a fixed load impedance). Thus, the selectivity of the filter may be changed by altering the coupling, a relatively simple matter with rotatable coupling loops.

Insertion Loss

Filter losses at resonance are primarily caused by dissipation and mismatch. Dissipation within the filter may be kept to a minimum by using a high-conductivity material on the inside surfaces of the cavity. A good, low-impedance short circuit at the short-circuited end of the line is also important. As in a lumped-element circuit, dissipation results in a finite unloaded Q. The magnitude of the dissipation loss generally lies between 0.25 and 2 decibels (db) and is given by[1]

$$\text{dissipation loss, db} = 20 \log \frac{Q_u}{Q_u - Q_l}$$

where Q_u = unloaded Q and Q_l = the loaded Q.

The magnitude of the mismatch loss is determined primarily by the relationship between input and output coupling factors (M_1 and M_2 in Fig. 2) and the load

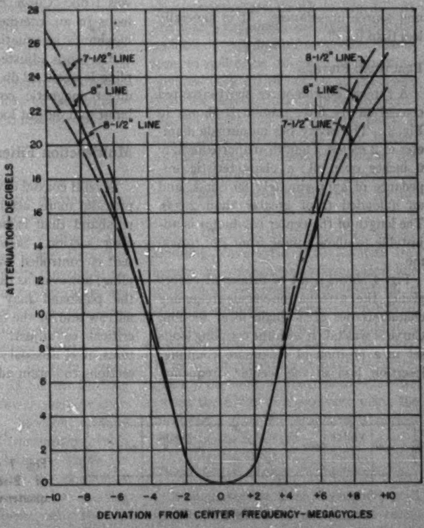


Fig. 5. Attenuation versus frequency characteristics of 2-section coaxial cavity filter for various lengths of coupling line between filter sections

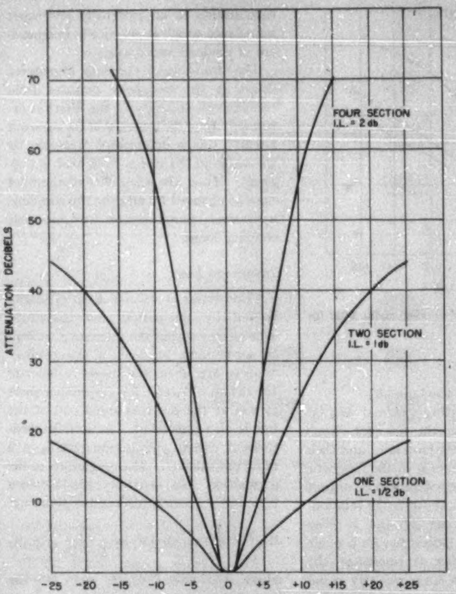


Fig. 6. Attenuation versus frequency characteristics for one, two, and four filter sections in tandem

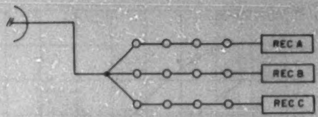


Fig. 8. Block diagram showing filter arrangement for multiplexing three receivers on common antenna

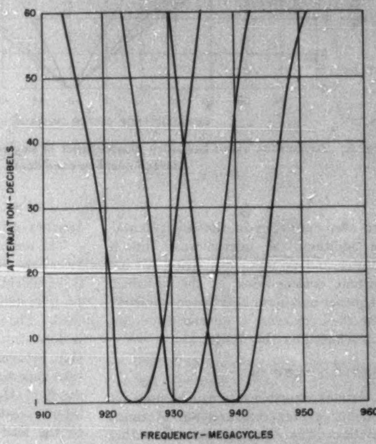


Fig. 9. Attenuation versus frequency characteristics of three flanking 4-section filters as used for receiver multiplexing

and source impedances. It is generally less than 0.1 db.

Practical Filters

A typical quarter-wave short-circuited coaxial cavity filter section is shown in Fig. 3. The cavity has an outside diameter of 2 inches, a length of approximately 6 inches over all, a characteristic impedance of approximately 80 ohms, and an unloaded *Q* of greater than 2,000. The length of the center conductor is adjustable to allow tuning from 890 to 960 mc.

Fig. 4 shows a set of measured curves giving the attenuation-versus-frequency characteristics of a single filter section. Curve *A* was taken with the coupling loops set in a position of maximum coupling; insertion loss at the center frequency was 0.25 db. Curve *C* was taken with the coupling loops set to allow very little coupling; insertion loss for this position was 1 db. Curve *B* was taken with the loops in an intermediate position. The maximum attenuation for the condition of the loops adjusted for maximum coupling is about 32 db, and is limited by the direct magnetic coupling between the input and output loops.

Multisection Filters

Several coaxial cavities are cascaded to provide more attenuation outside the passband than is possible with a single filter section. Selectivity to some extent is controlled by the degree of coupling; however, as the coupling is reduced, the passband may become intolerably narrow and/or the filter may become too critical to adjust. Under such conditions, it is necessary to add one or more sections to obtain adequate selectivity.

The most convenient method of coupling between filter sections is through transmission lines of an odd multiple of quarter waves in length.[9] This length is not critical, however, as shown in Fig. 5, where the attenuation characteristics are plotted for several lengths of coupling line.

Fig. 6 shows a set of curves measured with one, two, and four filter sections in tandem. Each section was adjusted to provide the same attenuation and insertion loss. The total insertion loss of the filter sections in tandem at resonance is the summation of the insertion loss of the individual filters; however, it is interesting to note that the attenuation for a given frequency outside the passband is greater than that of the sum of the attenuation of the individual filters. This is because of the increased mismatch loss created by the impedance-transforming properties of a quarter-wave length of transmission line. This can best be described by reference to Fig. 7. At a given frequency outside the passband, the input and output impedances of the filters

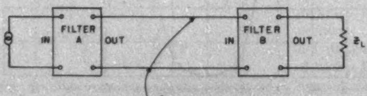


Fig. 7 (left). Block diagram of 2-section filter showing quarter-wave length coupling line

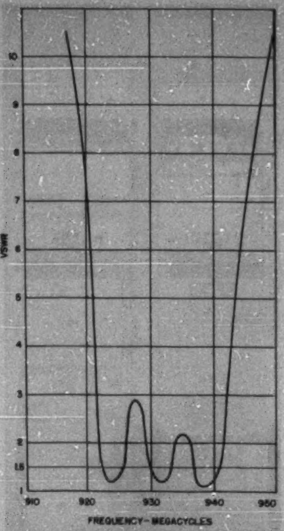


Fig. 10. Variation of VSWR with frequency on antenna transmission line. Three receivers multiplexed through 4-section cavity filters were used as load

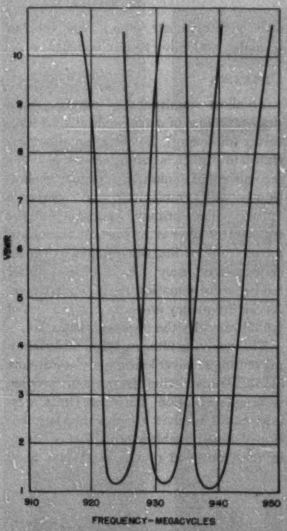


Fig. 11. Variation of VSWR with frequency on individual receiver lines using 4-section filters and antenna as load

are less than the characteristic impedance of the coupling line. Because the load impedance (Z_1) of the coupling line is less than its characteristic impedance (Z_0), the input impedance (Z_i) to the line will be greater than its characteristic impedance, i.e.

$$Z_i = Z_0^2/Z_1$$

With a zero-length coupling line (or a line a multiple of half-waves in length), the mismatch loss would be proportional to the ratio Z_0/Z_1; however, with a quarter-wave coupling line, the mismatch loss is proportional to Z_i/Z_1 which is equal to Z_0^2/Z_1^2. Loss at the passband frequencies is not affected, since the line is terminated in its characteristic impedance.

Applications

Coaxial cavity filters are used at present with Lenkurt type-*72B* radio systems for: 1. receiver pre-selection; 2. receiver multiplexing; 3. transmitter harmonic suppression; 4. transmitter multiplexing; 5. transmitter-receiver diplexing; and 6. multidiplexing.

Receiver Preselection

When the first stage of a superheterodyne receiver is a mixer, some form of filtering must be provided between the antenna and mixer to prevent extraneous signals from entering the mixer and causing interference. A 2-section filter having the attenuation characteristics shown in Fig. 6 is generally used on the type-*72B* receivers for preselection. In applications where extremely close spacing of transmit and receive frequencies is necessary, a 4-section filter may be used to provide the required additional selectivity.

Receiver Multiplexing

When used to multiplex more than one receiver on an antenna, the coaxial cavity filters must furnish sufficient isolation between receivers to prevent loading of the antenna at frequencies other than the operating frequency of the individual receivers. When coaxial cavity filters are used for multiplexing, they also act as receiver preselection filters, since the requirements for preselection are less severe than for multiplexing. In many cases, sufficient isolation could be provided by a 2-section filter between each receiver and the common line; however, because of the multiplicity of transmitters which usually exists at a location where receivers are multiplexed, the possibility exists of spurious frequencies' causing objectionable interference into the receiver. For this reason, the greater selectivity provided by 4-section filters is desirable.

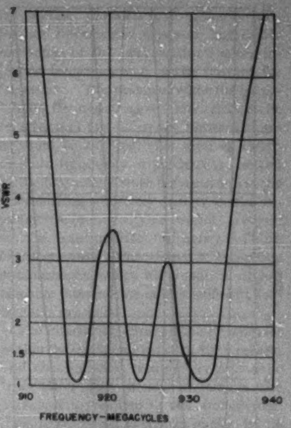


Fig. 12. Variation of VSWR with frequency on antenna transmission line. Three transmitters multiplexed through 2-section cavity filters were used as load

Fig. 8 shows a block diagram of the arrangement commonly used to multiplex more than one receiver on an antenna. The total insertion loss of each 4-section filter is 2 db. Attenuation-frequency characteristics are shown in Fig. 9. As previously explained, the filter input impedance is very low off resonance; therefore, to make the off-resonant input impedance very high (looking at the filters from the junction), the line interconnecting the filters and the junction block is made three-quarters of a wave length. At the resonant frequency, of course, the line input impedance at the junction is almost equal to the receiver input impedance.

Fig. 10 shows the variation of voltage standing wave ratio (VSWR) with frequency on the antenna transmission line with three receivers multiplexed on a common antenna. The maximum VSWR at resonance is 1.20, which results in a mismatch loss of less than 0.1 db. Fig. 11 shows the variation of VSWR with frequency on each of the receive lines between the filters and receivers, using the filters as a load and assuming a perfect match at the receiver.

Transmitter Harmonic Suppression

A single-section filter is generally used in the transmitter to provide the neces-

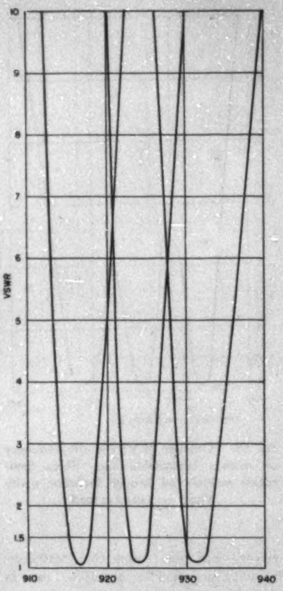


Fig. 13. Variation of VSWR with frequency on individual transmitter lines using 2-section filters and an antenna as load

sary suppression of harmonic radiation. A filter having attenuation characteristics as shown in curve *A* of Fig. 4, will generally ensure that harmonic radiation is at least 70 db below the unmodulated carrier.

TRANSMITTER MULTIPLEXING

Transmitter multiplexing filters must provide sufficient decoupling between transmitters to prevent the creation of intermodulation products of a disturbing nature in the final transmitter stage; in addition, they must allow loading of the antenna only at the operating frequency of each transmitter. Filters which provide the necessary attenuation for multiplexing automatically provide the required harmonic suppression, since the latter requirement is less severe.

It has been found that at least 25 db of isolation is required between type-*72B* transmitters to insure that spurious products will be more than 75 db below the transmitter unmodulated carrier. This requirement is particularly severe in the case where three parallel links are in operation with transmitters equally spaced in frequency. Under these conditions, the third order $(2F_c - F_h)$ or $(2F_c - F_l)$ product between the center- and high- (or center- and low-) frequency transmitters falls directly on the low- (or high-) frequency transmitter. This product, if it reaches the receiver with sufficient intensity, will result in an objectionable beat frequency which could cause interference in a carrier-derived voice channel.

A 2-section filter is adequate when the frequency separation between transmitters is at least 7.5 mc. However, where the separation is 6.3 mc, a 4-section filter must be used. Each filter section has an insertion loss of about 0.5 db and attenuation characteristics, as shown in Fig. 6.

Fig. 12 shows the VSWR variation with frequency of a 2-section filter on the common transmission line, using the filters as a load. Fig. 13 shows the VSWR characteristics of a 2-section filter on the individual transmitter lines, assuming a perfect match at the transmitter.

Figs. 10 and 11 show the VSWR characteristics for a 4-section filter. (Although the filters were discussed in connection with receiver multiplexing, the characteristics apply equally well for use with transmitters.)

Fig. 14 shows a photograph of two type-*72B* transmitters which will be multiplexed upon one common transmitting antenna. The multiplexer panel is shown near the top of the transmitter on the left-hand side; it has a capacity of three transmitters. Four-section filters are used in this case because the frequency separation between transmitters is only 6.3 mc. For this particular application, the traffic load is such that two circuits will provide the necessary capacity at this time; however, at the present rate of growth, it is estimated that within a few years one additional circuit will be required. The frequency for this additional circuit has been allocated and the multiplexer filters tuned and factory-tested. At the time of installation of the third transmitter, a cable patched between the multiplexer and transmitter power output panel will be the only requirement for connection of the new transmitter to the existing antenna system. Although not shown in the photograph, a similar arrangement is provided for the receiving system. This *72B* radio circuit has an ultimate capacity of 216 toll-quality telephone channels and 3 v-f order wire channels, plus a large number of alarm and pilot tone channels.

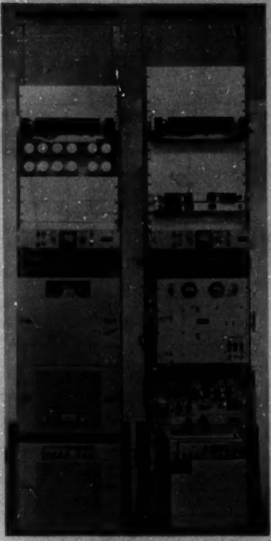


Fig. 14. Two type-72B transmitters multiplexed on common antenna. Multiplexer filter panel mounted near top of transmitter rack on left-hand side

DIPLEXING

To allow simultaneous operation on a single antenna of a transmitter and a receiver (diplexing), the filter arrangement shown in Fig. 15 is commonly used. The recommended minimum frequency separation between transmitter and receiver is 12 mc. The primary purpose of the transmitter filter is to prevent the transmitter from loading the antenna at the receiving frequency. A 2-section filter having an attenuation of 25 db at the receiving frequency and an insertion loss of 0.5 db provides the necessary isolation.

The receiving filter must, in addition to preventing receiver loading of the antenna at the transmitting frequency, provide sufficient isolation between the transmitter and receiver to prevent the creation of objectionable modulation products in the front end of the receiver. Such

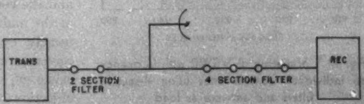


Fig. 15. Block diagram showing filter arrangement for diplexing

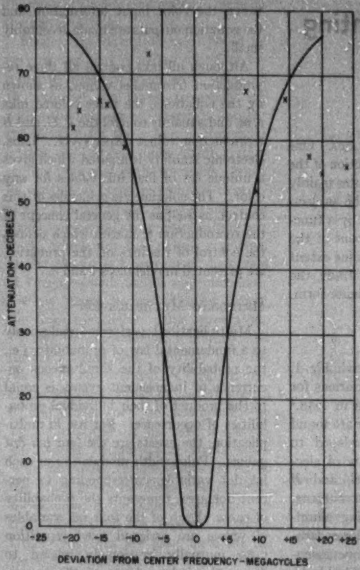


Fig. 16 (left). Attenuation versus frequency characteristics of 4-section filter. Crosses show required attenuation when diplexing

modulation products are primarily difference frequencies between high-order harmonics of the local oscillator mixer injection frequency and of the transmitter frequency. For example, when the transmitting frequency is lower than that of the receiver and the difference between them is equal to one-half the receiver intermediate frequency, a product is created in the first mixer stage between twice the transmitting frequency and twice the injection frequency. This product is equal to the intermediate frequency. There are numerous other discrete frequency combinations which will result in objectionable interference unless sufficient isolation is provided.

The interfering effect of the frequencies within the operating range of the receiver, and the isolation required between the transmitter and receiver determine the attenuation required to the transmitter. This requirement is shown in Fig. 16 on the same co-ordinates as a curve showing the attenuation provided by the 4-section receive filter. Fig. 17 shows a duplex transmitter-receiver assembly mounted on the same rack; the diplexer is mounted near the top of the rack immediately below the terminal block.

MULTIDIPLEXING

A frequently used combination of equipment, two transmitters and two receivers, may be operated on a single antenna through use of a multidiplexer; a combination of a multiplexer and a diplexer. The necessary filter characteristics for this arrangement may be obtained through analysis of the requirements for both the multiplexer and diplexer. Fig. 18 shows a photograph of two transmitters and two receivers which will be connected to a common antenna through the multidiplexer mounted near the top of the dual receiver rack. Four-section filters are used on the receiving lines and 2-section filters on the transmitting lines; the minimum frequency separation between a transmitter and a receiver is 12 mc; between receivers, 6.3 mc, and between transmitters, 7.5 mc. The spacing between transmitters could be reduced to 6.3 mc through use of 4-section filters on the transmitting line; however, this results in an unwieldy mechanical arrangement not conducive to convenient connection at the common junction block.

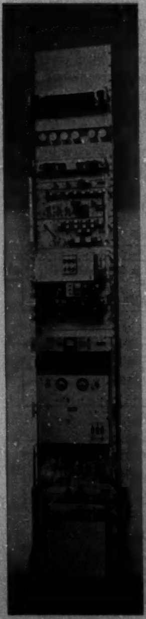


Fig. 17 (below left). Duplex transmitter-receiver assembly with diplexer mounted near top of rack

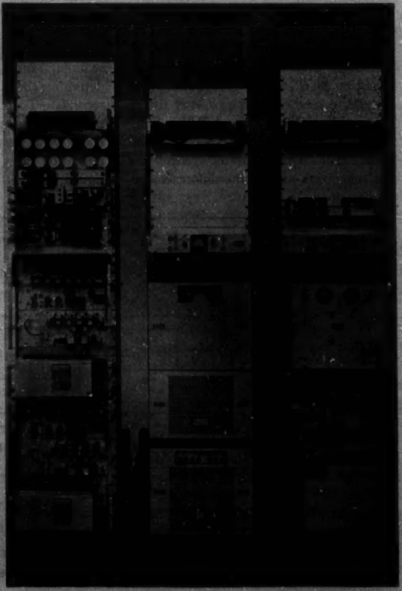


Fig. 18 (below right). Two transmitters and two receivers multidiplexed on common antenna. Filter panel near top of dual receiver rack on left-hand side

References

1. RADIO ENGINEERING (book), F. E. Terman. McGraw-Hill Book Company, Inc., New York, N. Y., 1947.

2. MICROWAVE TRANSMISSION CIRCUITS (book). Massachusetts Institute of Technology Radiation Laboratory Series, McGraw-Hill Book Company, Inc., New York, N. Y., 1948.

Electronic Computer for Color Printing

H. E. ROSE
NONMEMBER AIEE

PROBLEMS of color deficiency have been associated with color photography and magazine printing for many years and, more recently, with color television. The development of the additive and subtractive color theories in which the reproducing primaries are related to the spectral sensitivites of the 3-color-separation photographs, has been described by Hardy and Wurzburg, Jr.[1] A basic relation between the additive and subtractive contributions of the common printing process for a 3-color process was published in the same year.[2] The equations derived relate the transmissions of the 3-color-separation photographs to a function of the amounts of the inks and their tristimulus values. With the daily production of millions of magazine issues in the United States using a fourth (black) ink, compared to the smaller editions in Europe for which these equations apply, a modification to the equations was necessary to include the effect of the black ink used in the high-speed printing methods in this country. These final equations for color correction were published in 1948.[3] Where **V** is the vector of the red (*R*), green (*G*), and blue (*B*) tristimulus values of a color and the subscripts of **V** are the tristimulus values for the paper, full inks and full over prints of inks, the equation for the percentage of cyan (*c*), magenta (*m*) and yellow (*y*) inks is

$$\begin{aligned}\mathbf{V}=&(1-c)(1-m)(1-y)(1-n)\mathbf{V}_w+\\&c(1-m)(1-y)(1-n)\mathbf{V}_c+m(1-c)\times\\&(1-y)(1-n)\mathbf{V}_m+y(1-c)(1-m)\times\\&(1-n)\mathbf{V}_y+cm(1-y)(1-n)\mathbf{V}_{cm}+\\&cy(1-m)(1-n)\mathbf{V}_{cy}+my(1-c)\times\\&(1-n)\mathbf{V}_{my}+cmy(1-n)\mathbf{V}_{cmy}+n\mathbf{V}_n \quad (1)\end{aligned}$$

As simultaneous equations, with the ink percentages as implicit functions of the *R*, *G*, and *B*, color tristimulus values, they are as yet indeterminate with four unknowns and three equations. One unknown *n* can be made a function of the other three because it contributes mainly to the luminance control of the subject. Thus, with *n* being generated by a function of the luminance component of the other three inks, modified to some extent by preferences of the printing trade, the equations reduce to a determinate form.

Computer System

The computer system is shown in Fig. 1. The method of solving the equations for color correction was published in 1948.[4] Greater stability and accuracy are found in the present computer developed to solve faster for the higher rate of electronic scanning. Input *R*, *G*, and *B* signals have current generator outputs. They are derived by scanning simultaneously the *R*, *G*, and *B* color separations by a flying-spot technique essentially as described by Haynes.[5] The blocks marked *R′*, *G′*, and *B′* include the multiplication and summation of the right-hand side of the equations. Actually, the quadruple products of the ink values are derived and multiplied by three normally different values of subscripts of **V**. These are then summed in their corresponding *R′*, *G′*, and *B′* groups along with the $n\mathbf{V}_n$ terms. The *R′*, *G′*, and *B′* are compared to the *R*, *G*, and *B* signals by the e_r, e_g, and e_b units respectively and these error signal outputs control the ink generators so that the differences between the color input information and the equation outputs are made acceptably small.

Although all inks control all three reproduction tristimulus values, as shown by the equations, the three colored inks *c*, *m*, and *y* mainly control the *R*, *G*, and *B* reproduction values respectively. Thus, electronic stability is assured which gives a unique set of four ink values for any color. The colorimetric principles of this control, as well as the general concept of the reproduction primaries which defines the control of the inks on the primaries, are presented in references 1 and 3.

Method of Multiplication

Multiplication is performed analogously to a fundamental law of probability, i.e., the probability of the simultaneous occurrence of independent events is equal to the product of their individual probabilities of occurrence. For use in multiplication, the events are the four ink-dot values. Pulse-width modulation of each ink-dot variable, corresponding to percent dot area, represents the probability of each. Each of the four ink variables of pulses are assigned fixed repetition rates mutually irrationally related to insure the independence of the events. Just as in probability, where a large number of trials is necessary, the coincidence of the ink signals must be averaged over a long time. The time for scanning a unit area of picture must encompass many of the coincidence pulses. The scanning speed and size of the unit area determine this time to be about 100 microseconds. To resolve the area, the computer must be appreciably faster by being stable in at least 50 microseconds. The repetition rates of the four inks are assigned between 230 and 131.5 kc so that

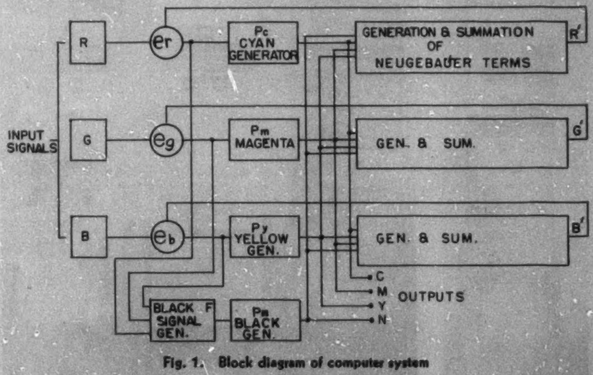


Fig. 1. Block diagram of computer system

Paper 55-180, recommended by the AIEE Computing Devices and Instruments and Measurements Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted August 5, 1954; made available for printing January 11, 1955.

H. E. Rose is with the Radio Corporation of America, Camden, N. J.

The author is indebted to C. B. Kinne, formerly of the Interchemical Corporation, for many suggestions during the initial development and to H. J. Woll, Radio Corporation of America, for suggestions on analytical approaches.

15 cycles of the lowest frequency occur in the unit-area time.

Ideally, the repetition rate should be very high so that in a fixed unit time for computer solution many more than the 15 cycles of the lowest frequency are within the unit time. However, another specification for the system performance is that a 1-per-cent pulse width, representing 1 per cent of the ink value, should be capable of being reproduced electrically. This determines the rise time of the signals supplying the multiplication networks which, for the highest repetition rate of 230 kc, is 1 per cent of the period of 4 microseconds or 0.04 microsecond.

Fig. 2 shows the multiplication circuit, the ink tristimulus value generators, *V1* and *V2*, the summation method, and the error signal unit. Starting with the ink dot *C* and reverse $(1-C)$ ink-dot generator, for the *C* values at the left, groups of four diodes each are shown receiving signals from the ink and reverse ink signal matrix of horizontal signal lines. Whereas all the reverse values are fed to the white group of diodes, the *C* group, for example, is the same except that *C* is used instead of the reverse *C* value. Other groupings are used until all the terms are available.

Each group of four diodes will then develop one of the products of the ink factors in the four color equations. The diode configuration will generate an output voltage at ground potential only when all inputs are at ground potential. Because the fraction of time that each signal is at ground is made equal to the ink-dot value, the average output value will be the product. Before filtering the output of each, which would give the average value required, the output of each group is fed to grids of three pentodes. These gate on to a current, assigned by the cathode resistor, which are the *R*, *G*, and *B* filter-emulsion constants for that ink. These output-current pulses are then summed by paralleling all plates for the appropriate color equation. These are then integrated in a single capacitor, one for each equation summation. Integration prior to this point would not allow the summing tubes to operate on an on-off condition. Thus, linearity demands on tube characteristics are avoided. Variations in tube emission are minimized by the heavy cathode degeneration developed by the controls for the *R*, *G*, and *B* constants.

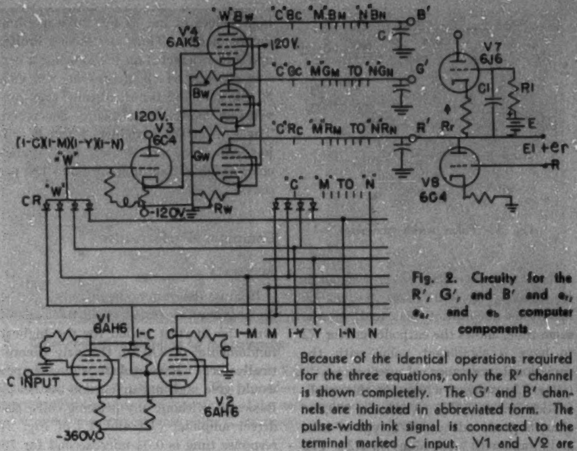


Fig. 2. Circuity for the R′, G′, and B′ and e_r, e_g, and e_b computer components

Because of the identical operations required for the three equations, only the R′ channel is shown completely. The G′ and B′ channels are indicated in abbreviated form. The pulse-width ink signal is connected to the terminal marked C input. V1 and V2 are push-pull power stages to drive the diode product generators "W," "C," and "M" to "N." The (1−M), M, (1−Y), Y, (1−N), and N signals are generated by stages identical to V1 and V2. Paper and ink color separation values are entered by the controls R_w, R_c of the stages similar to V4 to produce the equation terms WR_w, CR_c, etc.

PULSE-WIDTH SIGNAL GENERATORS, AMPLIFIERS, AND COMPARISON CIRCUITS

The conversion to a pulse-width signal from the error signal unit is accomplished by adding the signal voltage to a saw-tooth voltage wave form of fixed frequency and then referencing the sum with respect to the critical voltage of a trigger circuit. With the sum above the critical voltage, the trigger circuit is forced to one of its stable conditions. Below the critical voltage, the other stable condition occurs. A differential between the ON and OFF critical voltages is made to be somewhat less than 1 per cent of the saw-tooth height so that stable 2- and 98-per-cent pulse widths are available.

Fig. 3 shows a crystal-controlled oscillator synchronizing each of the saw-tooth generators to one of the four irrationally related frequencies required in the multiplication method. The high stability in the saw-tooth amplitude is mainly the result of the manner of operation of the half of tube *V4* marked "*B*". The output-wave amplitude is essentially equal to the difference in plate voltage between its two states; it is either cut off or gated on to a current under heavy cathode degenerative feedback. Capacitor *C* is charged to a fixed potential in a short time and discharges 25 volts through *R1*. *R2* and *R3* sum the input signal and the saw-tooth which is then referenced to the trigger circuit input at *V5* grid. Tubes *V5* and *V6* operate as a bistable circuit because of controlled, regenerative feedback to the screen of *V5*. The halves of *V6* are either conducting or nonconducting. Just as with the saw-tooth generator, this type of operation assures reliable levels of operation. The coupling circuit between *V3* and *V5* maintains adequate transient response and eliminates grid current effects on the saw-tooth wave form.

The circuits for the push-pull output stage of Fig. 2 and the d-c amplifier developing the amplified error signal are commonly used types of circuits. The amplifier features a gain of 50 and a −3-decibel response at 1 megacycle with a 2-tube subminiature cathode-coupled input stage followed by a cathode-coupled dual triode stage. Referred to the input of the amplifier, the drift is less than 2 millivolts over 1 hour after warm-up time. One hour is equivalent to 1 complete operation cycle of the apparatus.

The error signal is derived by providing a high variational resistance to the input- and output-current generators. Fig. 2 shows one of the circuits in detail for the *R* and e_r channels. With a signal of opposite polarity to *R* supplied to *V8* by an amplified multiplier phototube signal, the integrated current *R′* and $-R$ plate current of *V8* add to give a constant sum for perfect solution of the equation. An error signal output e_r will be generated for an inequality of the equation depending upon the variational resistance seen at the e_r signal line. This resistance is effectively increased by using *V7* as a

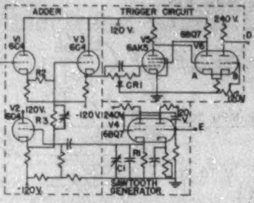


Fig. 3. Pulse width generator

cathode follower with the error signal as its floating ground. The battery *E* maintains a constant bias and, at the same time, allows the cathode resistor *R2* to have a significant value with the variational resistance of predominantly $(1+\mu)R2$ resulting, where μ is the amplification factor of *V7*. A value of 200,000 ohms is produced by a *6J6* type of tube as *V7*, and with 7,000 ohms for *R2*. The miniature, 45-volt battery operates at practically zero load.

BLACK-INK SIGNAL GENERATOR

The function required to generate a black-ink signal is shown in Fig. 4 where *C*, *m*, and *y* ink values are connected to two sets of limiters. The upper set limits the ink values individually to a value slightly in excess of maximum values of the three colored inks used with solid black ink to make the darkest color in the picture. High-speed printing usually assigns a total of 260 per cent of ink made up of 100 per cent black ink and 160 per cent divided almost equally among the three colored inks. Thus, ink values above approximately 53 per cent, which normally contribute towards saturation, are effectively eliminated as control factors in black-ink generation. The arithmetic adder 1 then provides a sum of ink values from which an amount may be subtracted equal to the amount of three-color black that may be desirable before black ink is to be allowed. The second limiter provides the solution to the problem that the last mentioned sum tends to generate as much black for two high values of ink as for three medium values of ink. Thus, the second limiter and adder detect if an ink is below a certain amount which, if present, will gate off the gate and threshold unit, thus preventing black from being generated. The gain of the amplifier *A* can be adjusted to give any desired total undercolor of ink within limits of 60 to 180 per cent.

Computer Performance

ACCURACY

Because the gated output pulses of the diode groups represent small fractions of even the shortest period for the highest fundamental of 230 kc, the only theoretically true rectangular wave shapes would give high multiplication accuracy. Basically, without any peaking coils, the driver amplifier (*V1* and *V2* of Fig. 2) response time is 0.04 microsecond for 70 per cent of full amplitude. A serious loss would occur, however, in the peak pulse output because of the exponentially changing wave form. Peaking coils, giving 5-per-cent overshoot, reduce product errors by a factor of 3 in the pulse-width region below 3 per cent. Products of three signals, two at fixed 50-per-cent pulse widths and the third of variable pulse width have been measured to be within 1-per-cent accuracy, with the greatest error occurring in the region below 3-per-cent pulse width of the variable signal.

Accuracy of computing is measured by the difference between the input color signals and the equation values for the ink colors. This difference must be held to a low value with respect to the *R*, *G* and *B* color separation values. Because the trichromatic coefficients are the ratio of the three filter-emulsion values, chromaticity values are independent of the brightness. Color accuracy is then most difficult for low-brightness colors, where the absolute values are in the order of 4 per cent. The specification of accuracy in this apparatus has been assigned to be such that a 1-per-cent difference in the trichromatic coefficient should be resolved. Then considering the low-brightness value of 4 per cent, the *R*, *G*, and *B* equations must be capable of being solved to an accuracy of 1 of 4 per cent or 0.04 per cent of their full range. The error signal outputs, described in the foregoing and derived by tube *V7* of Fig. 2, as well as the amplifier gain and duty-cycle generator will provide this performance.

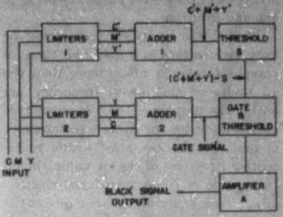


Fig. 4. Black-ink signal generator

A significant contribution towards accuracy is the extensive use of the on-off type of operation of tube circuits. This has been described as the basis of multiplication and of the pulse-width ink generators. The entire computer is supplied by electronic regulated power supplies, so that the remaining d-c amplifiers have a stable reference. The over-all computer accuracy is given by a tolerance of ±2.5 per cent in the ink values.

SIGNAL/NOISE RATIO

In any such pulse-width modulation information signal, a Fourier series specifies the fundamental and harmonic content. Four such pulse-width signals, modulating one another, will give rise to frequency terms of the fundamentals, beats of the fundamentals, and of all combinations of their harmonics. The Fourier coefficients reduce rapidly with the number of the signals and inversely proportional to the product of the harmonics of the individual frequencies. A single pulse-width-modulated signal is described by the equation

$$F_1(t)=c+\sum_{i=1}^{\infty}\frac{2\sin \pi ic}{\pi i}\sin iW_1t \tag{2}$$

where

c=fractional pulse width
i=harmonic

Multiplying four such functions of W_1, W_2, W_3, and W_4 frequencies with i, j, k, and 1 harmonics respectively gives the form of

$$F(t)=F_1(t)F_2(t)F_3(t)F_4(t)=cmyn+F(w) \tag{3}$$

Table I. Fourier Series Expansion of Modulation Products

Frequencies	General Expression of F(w)	Maximum Value
iW_1	$\frac{2myn\sin(\pi ic)}{\pi i}$	$\frac{0.637}{i}$
iW_1-jW_2	$\frac{2yn\sin(\pi ic)\sin(\pi jm)}{\pi^2 ij}$	$\frac{0.202}{ij}$
$iW_1-jW_2+kW_3$	$\frac{2n\sin(\pi ic)\sin(\pi jm)\sin(\pi ky)}{\pi^3 ijk}$	$\frac{0.0643}{ijk}$
$iW_1-jW_2+kW_3-lW_4$	$\frac{2\sin(\pi ic)\sin(\pi jm)\sin(\pi ky)\sin(\pi ln)}{\pi^4 ijkl}$	$\frac{0.0204}{ijkl}$

Vol.
Pt. 1
195

Fig. 5. Color correction supplied by the computer

Left: print of red separation negative. Center: print of the recorded, corrected C negative. Right: the computer correction mask is a unity-gamma photographic print of the composite of the first and the negative of the second. This procedure presents the computer's correction in a visual manner so that the density variations of the mask compared to the density of the lower step of the mask's gray scale determines the ratio of the corrected value to the uncorrected value of the C information

where $F(w)$ is a function of frequency, as shown in Table I, for all the frequencies and expressions of the terms for those frequencies.

The greatest amplitude occurs when, e.g., in the second row sin ic and sin jm both equal unity. This occurs when the harmonic factors i and j are equal to 1. The other two ink-dot value factors of c and m give the greatest amplitude when they are equal to 1/2.

Fortunately, because of the relatively low signal range of approximately 26 decibels for the process, it was found empirically that the coefficients below 0.02 may be disregarded. Where in row two, frequencies of beats of harmonics up to a product of ten are shown to contribute, row three indicates that harmonics only up to a product of four need be considered. These then, for the four frequencies used, show no more than 5-per-cent amplitude in the range of direct current to about 10 kc. Since these occur under unique values of colors, the average beat is of less than 5-per-cent amplitude. In fact, the peak-to-peak noise of the computer with the inks adjusted for maximum noise is less than 3 per cent or greater than 30 decibels down from the full signal. Although this signal-to-noise figure does not compare with those of most audio equipment, corrected plates recorded show no significant degradation of picture quality.

FREQUENCY RESPONSE

Speed of computation is established so that detail loss in the final photographic recording is not significantly attributable to the computer. As pointed out before, the time for scanning a unit area is 100 microseconds. Since a complete change of information may occupy this time, in order not to suffer detail loss the computer should become completely stable at both levels of the signal. A decision was made to provide aperiodic response to the input signal which is conveniently gained by having only a 6-decibel loss per octave in the computers frequency range. Thus, specifying that the response is to be within 2 per cent of its final solution for one-half the unit area time means that the −3-decibel response must be for a frequency of 13 kc. The 2 per cent of final solution figure was chosen so that the logarithmic recording on photographic materials of the transient response is not deteriorated in density by more than 10 per cent; i. e., although a voltage decreases to 2 per cent of its final value in 50 microseconds, the transient is recorded on the photographic plate to only 90 per cent of the full range of density available. It has been estimated that this loss is all that may be tolerated visually in order not to lose significantly on resolution.

The equivalent circuit of this computer can be shown to be simply the parallel combination of the two current generators R and R', the variational resistance R_r of the error signal unit ($V7$ resistance of Fig. 2), and the capacitance C of Fig. 2. An equivalent conductance K converts the error signal voltage e to produce the ink value which is returned to the R' generator. Also cross coupling exists between the m, y, and n ink generators to the R' generator. Therefore, the feedback loop contains a variable gain factor g which varies from near zero to near unity, depending upon the color region and the luminance value being computed. Neutral grays give the larger values of g which give a −3-decibel response frequency of 13 kc. This frequency may be shown to be

$$W_0' = KgRW_0$$

where

$$W_0 = 1/CR_r \qquad (4)$$

Fortunately, this high speed neutral gray and lower speed chromaticity computation is analogous to the sensitivity of the human eye. Just as in both the fields of color television, and color printing, the black or luminance information is considered a detail function or key plate. Although the conclusion may be drawn, therefore, that a recorded colored ink plate will not have the familiar photographic sharpness of detail in all colors, the 4-color printed proofs show acceptable visual resolution.

Conclusions

A number of conclusions may be presented from the description of the computer and from the experimental use in the printing trade.

1. An electronic computer has been described which solves three 4th-degree simultaneous equations in less than 50 microseconds to establish the four ink values of color printing required to match the color of the original picture.
2. A novel method of multiplication has

been analyzed for which simple circuit components perform with 1-per-cent accuracy.

3. Freedom from drift and accuracy is improved by the extensive use of the on-off, digital type of operation for vacuum tubes in the computer components. The computer accuracy will remain within ±2.5 per cent for a period of time greater than the cycle of operation required in producing the color-corrected plates of a subject.

4. Experimental use for a period of about 5 years has proved its usefulness because the parameters of the color input information system and the final processing steps of the plates required to make a printed proof have been considered.

5. Fig. 5 indicates the significant contribution of this system.

References

1. The Theory of Three-Color Reproduction, A. C. Hardy, F. L. Wurzburg, Jr. *Journal*, Optical Society of America, New York, N. Y., vol. 27, no. 7, July 1937, pp. 227–40.

2. Die Theoretischen Grundlagen des Mehrfachfarbenbuchdrucks, H. E. J. Neugebauer. *Zeitschrift für Wissenschaftliche Photographie, Photophysik und Photochemie*, Leipzig, Germany, vol. 36, no. 4, 1937, pp. 73–89.

3. Color Correction in Color Printing, A. C. Hardy, F. L. Wurzburg, Jr. *Journal*, Optical Society of America, New York, N. Y., vol. 38, no. 4, April 1948, pp. 300–07.

4. An Electronic Method for Solving Simultaneous Equations, A. C. Hardy, E. C. Dench. *Ibid.*, pp. 308–11.

5. The RCA-Interchemical All Electronic Color Correction System, Harold E. Haynes. *Penrose Annual*, London, England, vol. 46, 1952, pp. 83–86.

Open-Wire Carrier Systems in South Africa

C. F. BOYCE
NONMEMBER AIEE

THE UNION of South Africa and the neighboring mandated territory of South-West Africa occupy an area of about 800,000 square miles at the southern tip of Africa, about a quarter the area of the United States. It is a land of vast mineral wealth, and the world's chief source of gold and diamonds. Exceptionally rapid industrial development has characterized its cities during the past 20 years.

About a quarter of the total population of 14,000,000 live in the major cities and adjoining areas, shown in black in Fig. 1. The balance of the population is somewhat unevenly distributed throughout the country, and in towns, none of which have a greater population than 50,000. About one-fifth of the total population is of European descent.

The climate is in general a temperate one. A long range of mountains roughly parallel to the coast divides the coastal strip from the high plateau which forms a large portion of the total area of the country. Much of this plateau has a dry climate not unlike that of Southern California and New Mexico. The mean annual rainfall for South Africa is 25 inches, varying from 5 inches in the west to about 70 inches in parts of the eastern coastal zone.

There are nearly 600,000 telephones in South Africa, and of these just over two-thirds are dial telephones. All of the latter are in the nine major cities. The balance of the telephones are mainly of the local battery type employing 17-cycle-per-second (cps) signaling.

Within the areas shown in black in Fig. 1, telephone circuits are provided by underground cables, including carrier-on-cable and coaxial cable systems.[1,2] Outside of these areas all toll circuits are provided by open-wire routes and carrier systems on these routes. This paper deals with these open-wire toll telephone circuits and associated broadcast and telegraph circuits.

Growth and Present Size of Carrier Network

About 30 years ago, the first carrier systems, which were of the single- and 3-channel types, were installed in South Africa.[3,4] The subsequent growth in carrier telephone circuit mileage is illustrated in Fig. 2. The growth in carrier telegraph circuit mileage, i.e., voice frequency telegraphs employing carrier channels as bearer circuits, and carrier telegraphs (generally known as high-frequency or HF telegraphs) in the frequency range of 6 to 30 kc per second is shown in Fig. 3. There are about 600 carrier systems of various types in service, varying in length from 15 to 1,100 miles. The numbers of carrier toll telephone circuits on the main backbone routes only are shown in Fig. 1. In addition to these, carriers are extensively used on secondary routes.

The Scope of Open-Wire Carrier Systems

The minimum line length for which carrier exploitation appears to be economical on open wires depends on several interrelated factors such as the number of channels, the relative costs of the terminal equipment, the line, accommodation, availability of power and maintenance staff, etc. In some cases this minimum length may be as low as 10 to 15 miles. Frequently when a route cannot accommodate additional wires, the use of carrier systems is the only practicable method of increasing the circuit capacity. In general, for distances greater than 30 to 40 miles, and where the circuit growth is such that not more than about 200 circuits are likely to be required at the end of 20 or 25 years, an open-wire route with carrier systems may well prove to be the most economical solution, on a "present value of annual charge" basis.

The maximum number of telephone circuits at present provided on a single open-wire route is between 200 and 240. It is proposed to increase this figure in the near future, for distances up to 150 or 200 miles, to about 350 circuits, with the aid of compandors. This aspect is dealt with later in more detail.

Equipment Features

The principal features of open-wire carrier systems, many of which were first developed in the Bell System, are well known. Only those features which may differ from those employed elsewhere will be mentioned. The line and principal modulating frequencies of the systems used in South Africa are shown in Fig. 4. Several of these systems have been

---

Paper 55-231, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing December 1, 1954.

C. F. Boyce is with the Government Post Office, Pretoria, South Africa.

The author is indebted to the chief engineer of the General Post Office, Pretoria, South Africa, and to Standard Telephones and Cables Ltd., London, England, for permission to publish the information in this paper.

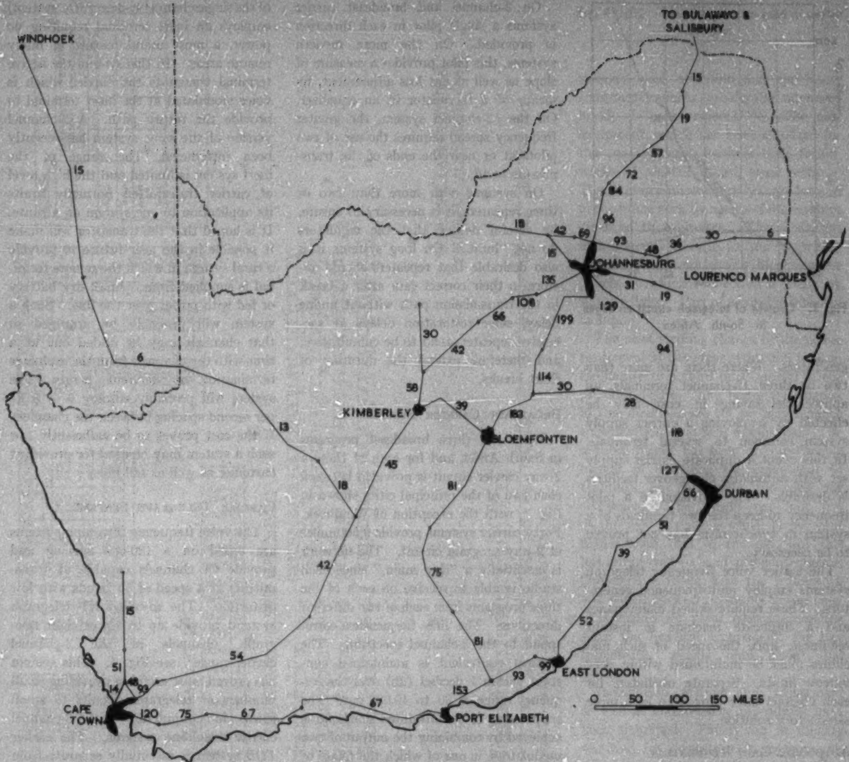

TO BULAWAYO & SALISBURY
WINDHOEK
15
6
13
18
45
42
54
15
51
48
14
93
CAPE TOWN
120
75
67
67
153
PORT ELIZABETH
81
75
81
KIMBERLEY
58
30
42
66
108
135
39
183
BLOEMFONTEIN
114
30
199
28
JOHANNESBURG
18
42
69
15
96
84
72
57
19
15
93
48
36
30
6
LOURENCO MARQUES
31
19
129
94
118
127
66
DURBAN
51
39
52
99
93
EAST LONDON
0
50
100
150 MILES

Fig. 1. Open-wire carrier telephone circuits in South Africa, on main lines only

Figures indicate the number of channels planned for 1958

described in various technical publications.[4-6]

Frequency Translation

Up to 1940 all systems employed modulation methods which translated individual channels direct to the required line frequency. Most modern systems employ one or more stages of group modulation, the advantages of which are well known. In the case of the *SOJ* 12-channel system, which is very similar to the *J* system of the American Telephone and Telegraph Company, the basic group of 12 channels is translated to the frequency range of 60 to 108 kc per second in one stage, employing crystal channel filters.[7] This base group is common to cable systems and much use is made of the facility to switch 12-channel groups without demodulating to voice frequencies.

Carrier Supplies

Single- and 3-channel systems employ inductance-capacitance oscillators individual to each channel. In practice these are found to be sufficiently stable to maintain channel synchronization to within 2 or 3 cps over periods of several months, particularly the more modern temperature compensated types. Individual channel oscillators render equipment more flexible, but in the case of a 12-channel system there are economic and technical advantages in employing a common crystal-controlled master oscillator in conjunction with harmonic

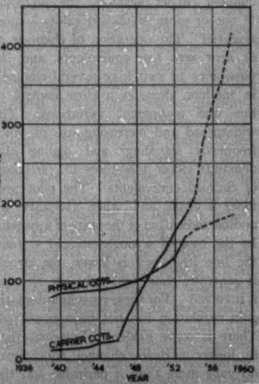


Fig. 2. Growth of toll circuits in South Africa: trunk circuit mileages

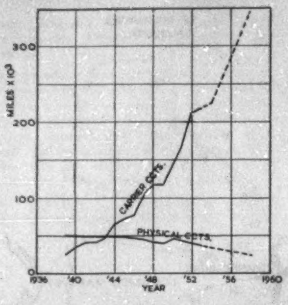


Fig. 3. Growth of telegraph circuit mileages in South Africa

generators. Where there are more than two or three 12-channel terminals, an appreciable saving in cost may be effected by employing a carrier supply system common to several terminals. In this event a duplicate carrier supply set with automatic changeover facilities is desirable. The provision of a pilot frequency to keep the two terminals of a system in synchronism has not proved to be necessary.

The earlier voice frequency telegraph systems employ multifrequency generators. These require skilled maintenance and a duplicate machine in instant readiness, since the speed of such machines must be maintained within very narrow limits. Separate oscillators for each channel have proved to be a more satisfactory solution.

Automatic Gain Regulation

With the possible exception of single-channel carrier systems, the provision of one or more regulating pilots is most essential. The operation of channels at or near zero loss equivalents, and the possibility of switching several channels in tandem, demand a close control of line and equipment variations. The widespread and increasing use of voice frequency signaling[9] and dialing also demands closer tolerances in variations of channel equivalents. The pilot provides a visual and, when necessary, an audible indication of the state of a system. In the case of a single-channel system, the line frequency is lower, the circuit is generally short and only one channel is involved. For these reasons automatic gain regulation has not been employed. There is some evidence, however, that in longer systems channel equivalents do vary more than is desirable and the provision of a pilot would also provide a system failure alarm.

On 3-channel and broadcast carrier systems a single pilot in each direction is provided. On the more modern systems, this pilot provides a measure of slope as well as flat loss adjustment, by means of a thermistor in an equalizer. On the 12-channel system, the greater frequency spread requires the use of two pilots at or near the ends of the transmission band.

On systems with more than two or three repeaters, it is necessary to ensure, by careful design, that the regulators do not "hunt." On long systems it is also desirable that repeaters should restore to their correct gain after a break in the transmission path without undue delay, since restoration delays at successive repeaters tend to be cumultative, and therefore extend the duration of short breaks.

Broadcast Carrier Systems

There are three broadcast programs in South Africa and for each of these a 2-way carrier circuit is provided between each two of the principal cities shown in Fig. 1, with the exception of Windhoek. Forty carrier systems provide 9,500 miles of 2-way program circuit. The network is essentially a "ring main," since each studio is able to receive on each of the three programs from each of two different directions. The line frequencies correspond to the 3-channel spectrum. The channel equivalent is maintained constant within 1 decibel (db) over the frequency range of 30 to 10,000 cps. The suppression of the unwanted side band is achieved by combining the output of two modulators, in one of which the phase of both the program input and the carrier has been displaced by 90 degrees relative to the other modulator.[4] Both the noise and the distortion products of the channel are of a very low order. At intermediate studios which normally accept, and only occasionally originate, programs, the program is received "in tee" from a repeater to avoid introducing the effects of local cables. Facilities are provided so that the studio can at any time convert the repeater into two terminals and so originate a program.

Rural Carrier Systems

Apart from the main routes, carrier systems have played a vital role in providing toll circuits in rural areas. Although farms tend to be large and the population density low, most areas require telephone service at intervals of the order of 10 miles or less. Single- and 3-channel systems are extensively used to provide toll lines. One version of the single channel system (*SOC* system) employs an inert terminal requiring no power, a most useful feature in many remote areas. In this system the active terminal transmits the carrier which is voice modulated at the inert terminal to provide the return path. A 2-channel version of the same system has recently been introduced. The range of the inert system is limited and the high level of carrier transmitted normally limits its application to one system on a route. It is hoped that the transistor will make it possible in the near future to provide a rural system in which the remote terminal is supplied from a small dry battery or fed with power over the line. Such a system will probably be arranged so that channels may be added one at a time with the channels from the exchange terminating at different points. The system will probably employ 6 or 8 kc per second spacing between the channels. If the cost proves to be sufficiently low such a system may be used for providing customer as well as toll lines.

Carrier Telegraph Systems

The voice frequency telegraph systems are based on a 120-cps spacing and provide 24 channels capable of transmitting at a speed of 50 bauds with low distortion. The so-called HF telegraph systems provide up to six or eight telegraph channels in the 3-channel carrier range; see Fig. 4. This system has proved most useful in providing small numbers of telegraph circuits to small towns in conjunction with 3-channel carrier telephone systems. The earlier *TOB* system is essentially separate from the 3-channel telephone system, whereas the later *STO* type employs some equipment which is common to the *STO* telephone system.

All the systems, which are of the amplitude-modulated type, use rectifier modulators at the transmitting end and telegraph relays at the receiving end. Trials on frequency-modulated telegraph systems indicate that they are considerably less affected by short sudden changes of level. Frequency-modulated systems are, however, considerably more costly, bulkier and somewhat more complicated, and their use on stable high-quality bearer channels has not, up to the present, been justified.

Equipment Practices

The design of carrier equipment in regard to the quality of components, methods of mounting and wiring, long-term stability, ease of maintenance, and the saving of space has for many years

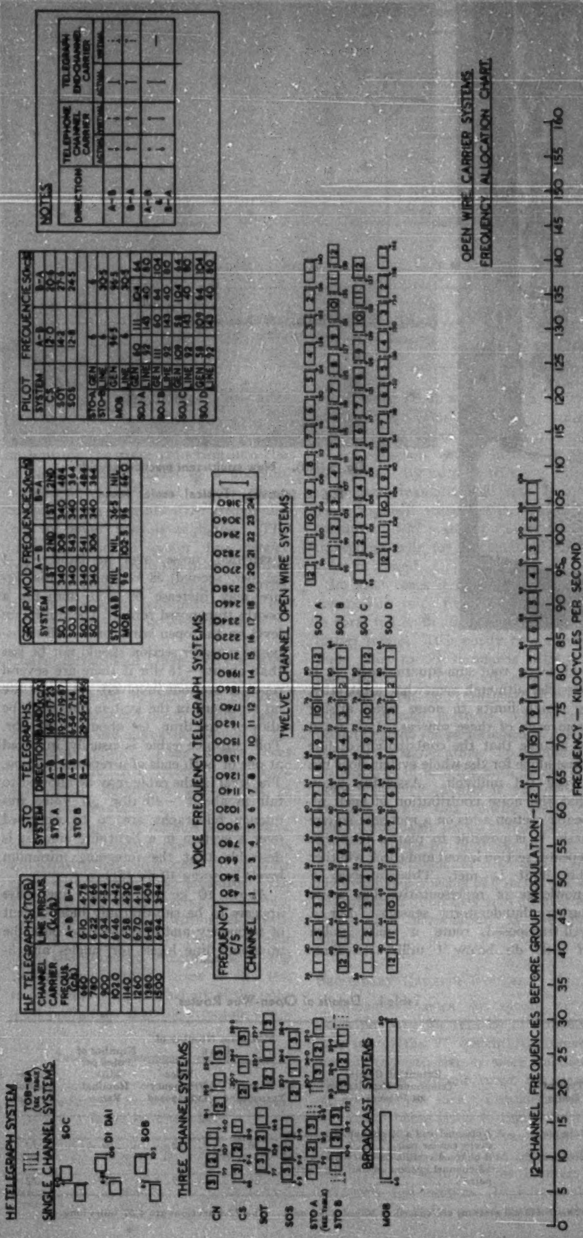


Fig. 4 (left). Line frequencies used in carrier systems

received close attention and the latest equipment reflects considerable advances in all of these respects.[10,11] The new equipment, which has been standard for the past 6 years, features light frameworks of welded folded steel sections. The slide-in equipment panels are located in the framework by means of die castings attached to each side. These castings mount terminal blocks for the bay wiring which fall opposite similar blocks in the panels, to which connection is made by suitable links. Part of a bayside is shown in Fig. 5.

The most striking change in the newer equipment is the great reduction in size, frequently to a third or a quarter of the previous figure. The limits to progress in this direction are set by maintenance considerations and heat dissipation. The reduced height has been of considerable benefit in installing the equipment in buildings not originally intended for equipment. The weight per foot run of the new bayside is some 30 per cent higher than that of the old, but this has not proved to be of any consequence. Shipment is made with the framework and panels packed separately in one case. This makes for great ease in handling since the heaviest component, the bay framework, weighs only about 100 pounds.

Installation testing and adjustment as well as maintenance procedures have been simplified. Panels can be quickly removed for inspection or replacement. All units requiring adjustment are arranged on the front of the panel. In addition to testing points on the bay-panel links, at which both terminated and level measurements may be made, important points are brought out to auxiliary terminal strips on the face of the panel, thus giving access for measurement without unsoldering wires. A centralized valve failure alarm panel enables space currents to be measured with facility and also provides a continuous monitor on space currents. Should any valve space current fall below a preset value, an alarm is operated.

Components may be sealed either individually and then mounted in a dust proof can, or alternatively assembled without individual seals, in hermetically sealed cans filled with dry air. The latter procedure is frequently cheaper in manufacture but does necessitate arrangements at central depots for the opening and repair of cans under skilled supervision. Experience has shown that

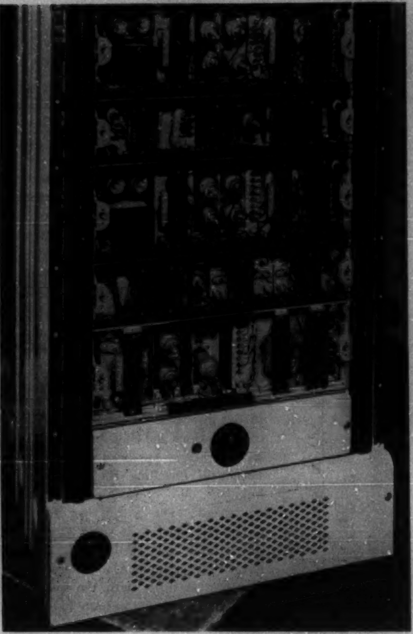


Fig. 5 (left). New equipment practice bayside

Fig. 6 (above). Typical carrier route

Vol. 7
Pt. 1
1955

the percentage of faulty cans is low, and faults occur mainly during installation. In general, only cans which contain critical components likely to be affected by humidity variations should be sealed. All equipment is wired either with polyethylene-insulated wire or where permissible polyvinyl-chloride-insulated wire.

The Siting of Carrier Repeater Stations

Since single- and 3-channel systems today rarely involve a repeater, the siting of repeater stations is mainly associated with 12-channel repeaters. The maximum line-transmitting level per channel is +17 db relative to the switchboard test level (dbr). The minimum level is set by noise considerations. A maximum noise voltage of 1 millivolt as measured on a European International Consultative Committee on Telephony psophometer and telephone weighting network measured at the 2-wire end of the circuit is permitted. This noise is made up of equipment noise, crosstalk from other systems on the route, and atmospheric or static noise. On the assumption that the various noise levels add on a root-sum-square basis, it is possible, although somewhat arbitrarily, to assign limits to noise levels arising from each of these sources. It is usual to assume that the contribution of line noise alone for the whole system does not exceed 0.3 millivolt. Assuming again that the noise contribution of each repeater section adds on a root-sum-square basis, it is possible to plan a tentative repeater section layout and check whether this limit is met. This involves a knowledge of representative line noise during thunderstorm seasons. For a well-transposed route a noise value of −85 db below 1 milliwatt (dbm) (weighted noise over a band of 4 kc per second) is assumed. This figure may increase to −70 dbm for a poorly transposed route. The minimum level on the open wires on a well-transposed repeater section should not be less than about −15 dbr if there are several repeater sections involved. If there are no repeaters in the system this may be allowed to drop to about −20 dbr. Toll entrance cable is usually involved at one or both ends of a repeater section. The level in the cable may be allowed to fall to about −40 dbr. If voice frequency telegraphs are to be operated over a system in a lightning area, it is desirable that the foregoing minimum levels be some 10 db higher.

Above 10 kc per second, line noise appears to be substantially independent of frequency and it is usual to base the minimum line levels for single- and 3-

Table I. Details of Open-Wire Routes

Type of Route	Extent of Carrier Telephone Exploitation at Present	Maximum Number of Pairs		Number of Poles per Mile, Nominal Value	Transposition System
		Carrier Transposed	Voice Frequency Transposed		
Main line	A 3-channel and a 12-channel system on every carrier pair	16	12	40	*J3*
Secondary	A 3-channel system on every pair and a 12-channel system on half the carrier pairs	16	12	28*	*J2*

* The fractional sections are spaced at 40 poles per mile. The *J2A* sections are 4.57 miles long.

Table II. Wire Configuration, Open-Wire Main-Line and Secondary Routes

Types of Pairs	Pairs per Arm	Spacing Between Wires of a Pair, Inches	Distance Between Centers of Horizontally Adjacent Pairs,* Inches	Distance Between Arms, Inches
Carrier pairs, to 1953	4	8	23	24
Carrier pairs on 40-pole-per-mile routes after 1953	4	6	26	24
Voice-frequency pairs	6	6	13	12†

* Pairs on either side of the pole are spaced 24, 23, and 23 inches respectively.
† The voice frequency arms are always below the carrier arms and 24 inches from the nearest carrier arm.

Table III. Fine Weather Copper-Wire Attenuation per Mile for 8-Inch Spacing, Db

Diameter, Inches	Pounds per Mile	10 Kc	30 Kc	90 Kc	150 Kc	270 Kc
0.137	300	0.070	0.120	0.208	0.271	0.34
0.112	200	0.087	0.144	0.245	0.317	0.40

channel systems on the same figures as given. The separation between open-wire telephone and power routes is based on limiting the noise induction into the physical or audio circuit to 1 millivolt (weighted value)[12] and under these conditions the noise at carrier frequencies due to power lines is negligible. The increasing use of power line carrier systems has not up to the present caused any interference to telephone carrier systems, but the possibility of its occurring should not be overlooked.[12]

Characteristics of Open-Wire Routes

The principal physical and electrical details of standard South African open-wire routes are given in Tables I, II, III, and IV.[4,8,12] Their design owes much to the Bell System[14] but there are certain differences arising out of local circumstances.

Since supplies of suitable local timber are not yet adequate, most carrier routes are constructed with steel or concrete poles with tubular steel crossarms. To maintain the wire spacing accurately, the arms are braced with steel straps; see Fig. 6. The arm (7 feet) is shorter than the standard American arm and the distance between arms is also less in order to reduce the height of the pole. This has meant that the crosstalk couplings are some 6 db worse than in American practice. This has been compensated on main lines by the use of the *J3* transposition system instead of the *J2* system. By keeping wire sag differences and irregularities in span lengths down to small values, it is possible to operate an *SOJ* system on every carrier transposed pair for distances of 1,000 miles or more.[3]

Table IV. Maximum Crosstalk

Worst Crosstalk Between Carrier-Transposed Pairs per Repeater Section Over the Frequency Range Indicated, Db

Type	10 to 150 Kc	150 to 250 Kc
Far-end	55 to 70	35 to 45
Near-end	40 to 65	30 to 45

In most cases it is necessary to accommodate voice frequency pairs, such as rural toll or multiparty farm lines on carrier routes. It is costly to construct these pairs to *J3* standards with four pairs per arm. These are now erected on 12-wire arms and transposed to simple voice frequency patterns. These pairs do not degrade the performance of carrier pairs on the higher arms. To standardize on one type of arm for both carrier and voice frequency circuits it was decided in 1953 to use this 12-wire voice frequency arm as an 8-wire arm on 40-pole-per-mile routes by omitting the center pair on each side of the arm. This involves a pair spacing of 6 inches and an improvement of about 4 to 6 db in crosstalk. The attenuation is increased slightly.

Secondary Carrier Routes

Secondary routes are spaced at 28 poles per mile for reasons of economy. The use of the *J2* system is, however, not entirely satisfactory since fractional transposition sections must be spaced at 40 poles per mile. Consideration is at present being given to the use of the *LJ* transposition system (a variant of the *J3* system) at 28 poles per mile for all types of section. This involves a 9.14-mile full section. It is probable that such a route would accommodate a 12-channel system on at least half the carrier pairs for three or four repeater sections. This would meet the requirements for most secondary routes. The use of 6-inch spaced pairs on 28-pole-per-mile routes for carrier working is not favored on account of the possibility of wind contacts. The use of copper-steel wire on such routes would enable the 6-inch spaced arm to be used. The high attenuation of such wire at voice frequencies limits its use to short lengths unless a satisfactory and inexpensive solution to the problem of providing gain in physical circuits is forthcoming. The possibilities of the transistorized negative impedance repeater in this connection are being explored.

Extension of Frequency Range Beyond 143 Kc per Second

The formation of ice on open-wire conductors is a rare occurrence in South Africa, and then only in very restricted areas. Thus the possibility of using frequencies appreciably above the *SOJ* range appeared to be restricted by crosstalk considerations. Crosstalk could be reduced further only by shorter transposition intervals, since crosstalk coefficients and irregularity crosstalk had been reduced as far as practicable. It is not economical to reduce span lengths further; at 40 poles per mile most routes are already stronger than they need be. The use of suspended transpositions, while valuable for river crossings, does involve a number of construction and maintenance problems.

Measurements indicate that six pairs out of the eight on the first two arms of a *J3* 40-pole-per-mile route have satisfactory attenuation characteristics up to at least 250 kc per second, while the worst far-end crosstalk, Table IV, has deteriorated by not more than 20 db. The availability of a compact and

Table V. One-Year Analysis of 12-Channel Network Outage

Outage per 100 Channel Miles per Annum, 21 Hours 24 Minutes

Cause of Fault	Percentage of Fault Time
Line failures	41.0
Equipment failures	
Consumable items	7.7
Nonconsumable* items	1.2
Adjustments	5.7
Power plant failures	18.2
Tests made outside of routine times	13.1
Faults which either came clear while being localized or where no faults were found	13.1

* Items such as condensers, rectifiers, etc., i.e., other than valves and fuses.

inexpensive compandor, giving a noise advantage of some 20 db or more, has thus made the application of another 12-channel group above the present *SOJ* systems on at least half the pairs a practical possibility. The top frequency of the new system is expected to be about 270 kc per second. The problems associated with such a project, e.g., method of frequency translation, pilot regulation, interaction crosstalk at repeaters, matching of entrance cables, etc., have been examined and it is proposed to install a number of systems in the near future. The same repeater sections will be employed since the compandor will be able to cope with the increased noise. It is proposed to employ these systems initially over not more than two or three repeater sections.

The Application of Carrier Systems to Early Types of Route

It is possible to operate one carrier system, even a 12-channel system, over almost any type of route, by suitably transposing the pair which will accommodate the carrier. The operation of more than one carrier system may require the insertion of a few poling transpositions or extensive retransposing.[8] It is also necessary to remove any sections of covered wire or cable interposed in the route.

Many of the earlier routes in South Africa embodied "revolved wires," i.e., a quad of four wires twisted or revolved making a complete twist in each four spans. Like all routes which employ special configurations to reduce mutual coupling, some pair combinations give good results while others are very poor. By a careful choice of pairs and the insertion of poling transpositions it is possible to operate several carrier systems over such routes. In one case 60 carrier channels were operated over such a route for 56 miles, although with somewhat relaxed crosstalk standards.[8] There is little doubt, however, that the flat carrier transposed route is superior to all other types of construction for carrier system operation.

Lead-in Arrangements

The *SOJ* carrier system with its extended frequency range introduces problems of matching between lead-in cables and the open wires not present with systems operating up to 30 kc per second.[7,8,15,16] The open-wire pairs are connected to the *SOJ* line filters by means of low-capacity star-quad cable having a characteristic impedance of 240 ohms. Variable load units are fitted to one or both ends of this cable and its impedance is built out to match the 575-ohm line, with a reflection coefficient of less than 5 per cent over the frequency range of 10 to 150 kc per second. In the frequency range of over 150 kc per second, it will be possible to permit a somewhat higher reflection coefficient, since the near-end crosstalk has not deteriorated as much as the far-end, the relative improvement being of the order of 10 db.

This arrangement is satisfactory for distances between the line and the equipment up to about 200 yards. For greater lengths it would be possible, but uneconomical, to use intermediate loading. In these cases the 12-channel line filters are divorced from the equipment and placed in a hut near the end of the route. The filters are connected to the terminal or repeater equipment by a paper-insulated star-quad 46-pair 16-gauge (40 pounds per mile) toll entrance cable. The high-pass side is brought in over an unloaded pair, giving a loss of 3.6 db per mile at 140 kc per second. The low-pass side is brought in over a pair loaded at intervals of 250 yards with 3-millihenry coils, giving a loss of about 1 db per mile at 30 kc per second. If the cable is short, the 3-millihenry loading may be dispensed with and autotransformers used at both ends of the pair for matching purposes. This latter arrangement does, however, introduce a considerable loss into the circuit and it is not favored for cables longer than about 500 yards.

In cases where only a limited number of 12-channel systems is proposed, and it is therefore possible to relax the impedance matching requirements, matching transformers covering the frequency range 0.5 to 150 kc per second are available. This transformer matches the open-wire line to the cable and eliminates the need for the filters at the terminal of the open-wire route.

Crosstalk in toll entrance cables is reduced to satisfactory values by suitably "doping" the joints on a basis of capacity unbalance and also by fitting crosstalk balancing condensers at one end of the cable.

Power Plant

Equipment installed up to about 1945 is operated from 24-volt and 130-volt d-c supplies. At most stations this is provided by one or more selenium rectifier units. The voltage regulation is achieved by tap-changing transformers controlled by what is essentially a moving-coil contact voltmeter. In the event of a failure of the public mains, or of the rectifier units themselves, the 24- and the 130-volt supplies are obtained from two d-c generators driven by an automatic-starting diesel engine. The break of about 10 seconds during which the engine is running up to speed is covered by very small batteries floated across the output of the rectifier units. On restoration of the mains supply the stand-by set recharges its starter battery.[6,8]

During the past 10 years, most carrier equipment has been operated from small power panels associated with each bay-side. Filaments are heated from a 6.3-volt a-c supply, while the plate voltage is 220 volts. This new arrangement has a great deal to recommend it,[1] though there are two important complications. In the first place, the mains voltage must be regulated to within 1 per cent and surges on the mains must be reduced as much as possible since they are applied directly to the equipment. Secondly, the problem of supplying a no-break stand-by power supply is not easily solved.

At less important stations, or where voice frequency telegraphs or broadcast carriers are not involved, automatic-starting Diesel-alternator sets are provided. In these cases there is a 5- or 10-second break in the supply on a mains failure. At all other stations a continuous a-c supply is provided by either of the following types of machine:

1. A synchronous motor running off the mains drives a large flywheel. On mains failure this motor supplies the load as an alternator driven by the flywheel and, at the same time, the flywheel starts the Diesel engine through a magnetic clutch.

2. The station is supplied from an induction-motor-alternator set running off the mains. This set drives a large flywheel. On mains failure the flywheel continues to supply the station by driving the set while the engine is brought up to near full speed before being magnetically coupled to the rest of the set.

There are several advantages in the second set described, e.g., mains voltage fluctuations tend to be smoothed out, the control circuits are simpler, and no synchronization of the machine with the restored mains supply is required.

The control equipment has proved to be the most vulnerable part of stand-by sets. It is essential that all relays and contactors be of the most reliable type and that all connections be as secure as possible. The sets should also be fitted with adequate and reliable alarms to cover a failure of any of the essential functions of the set.

Maintenance

Although maintenance standards on carrier systems have always been high, the tolerances of fault incidence and the stability requirements of toll circuits are even more stringent today, because of the following:

1. The automatization of the toll network in the form of operator dialing and in some measure customer dialing.[16]
2. The operation of circuits at or near zero loss equivalents, and the connection of several of these circuits in tandem.
3. The fact that a fault may affect a large number of channels.

EQUIPMENT PRACTICES

The effect of equipment practices on maintenance has already been mentioned. The more important points are:

1. Jack-in panels give ready access to components and wiring, and facilitate the repair and replacement of faulty panels.
2. The miniaturization of equipment should not be carried too far. The cost of floor space is low in relation to the equipment, and there are diminishing returns in miniaturization which may be absorbed by higher maintenance costs. It is, of course, also true that in some cases floor space is acutely limited, and in these cases very compact equipment may be essential.
3. There should be an adequate number of test points to facilitate the speedy location of faults. Where necessary it should be possible to make terminated, as opposed to level, measurements.
4. Controls, such as potentiometers, etc., should be concealed to prevent unwarranted or accidental operation.
5. The equipment should be reasonably dustproof.

RELIABILITY OF COMPONENTS

A system is no more reliable than its weakest component. Electrolytic condensers should in general be avoided. The voltage rating of condensers in decoupling circuits and in line filters must be generous. Wire-wound resistors should be used only for low-resistance values and where a high accuracy is necessary. Rectifiers in the form of metal rectifiers or germanium diodes are used extensively in modern circuitry and it is important that they should have stable characteristics. Catastrophic and emissive failures in thermionic tubes have been, and are still, receiving close and considerably successful investigation by manufacturers. Such special, or "trustworthy," tubes are of course more expensive than commercial tubes, but mean lives of between 3 and 10 years may be expected from them.

DEFECTIVE CONNECTIONS

As important as the components, are the connections between them. Careful control of soldering in both factory and field is essential. There is much to be said for soldering in tubes, particularly if they are reliable ones. Potentiometers employing contacts sliding over wire or studs, keys, jacks, blind sockets, riveted joints in links, or tags should be avoided. Flat-bladed links in spring sockets are superior to the round type. Contacts should always be of precious metal. The internal connections in such items as paper condensers, miniature rectifiers, etc., are possible weaknesses.

It is now the practice to subject all components and wiring after equipment installation to a vibration or percussion test.[17] Briefly, the method involves the sending of a test tone, of a suitable frequency and very low level, through the equipment to a demodulator followed by a high gain loudspeaker-amplifier. The disturbance of a bad connection modulates the test tone and produces a crackle from the loudspeaker. It is found that this method is very successful in revealing defective connections which appeared perfect under visual inspection. It is necessary to employ conscientious personnel for this type of testing.

TESTING EQUIPMENT AND MAINTENANCE AIDS

Terminal stations of any appreciable size and the more important repeater stations, are equipped with the following items of test equipment:

1. A continuously variable frequency oscillator covering the entire frequency range of the equipment, with the exception of the group translation stages in the *SOJ* systems. At larger stations there are additional oscillators, some covering the audio-frequency range only.
2. A nonfrequency selective transmission measuring set or sets to cover the entire frequency range of the equipment and capable of reading down to −40 dbm or in some cases—55 dbm. Most sets fall into one of the ranges, 30 cycles to 10 kc per second, 30 cycles to 150 kc, or 60 to 600 kc per second.
3. A multirange general-purpose 20,000-ohm-per-volt voltmeter with current and resistance scales.

At stations which have 3-channel terminals only, an 800-cps oscillator, together with a suitable level measuring set, is generally adequate. With some exceptions, no testing equipment is supplied for stations with single channel terminals only, since these are often unattended and small portable sets are carried by visiting maintenance personnel.

At the largest stations special items such as impedance bridges, psophometers, oscilloscopes, valve voltmeters, etc., are held. Harmonic analyzers for measuring the harmonic content of line amplifiers are also available at a number of stations but these have not proved convenient to use. Most equipment is mounted on suitable trolleys. This arrangement has proved most satisfactory since one set serves the maximum amount of equipment, the test leads are kept short, and the testing officer is in general conveniently near the equipment he is testing.

Experience has shown that it is important that station alarm systems be adequate, well maintained, and standardized throughout the country. The terminals and repeaters of important systems or routes are connected by a private telephone circuit, generally a physical line. On the longer systems a special circuit provides an immediate indication at both terminals of which repeater or repeater section has failed. The action of this circuit is quite independent of the faulty station or section since it is actuated by pilot alarms at adjacent repeaters. This device is of particular value in locating short breaks which have not been noticed by repeater attendants, and for calling out the correct attendant when faults occur during the night. It is important that the correct maintenance tools be available, particularly in regard to soldering equipment.

ROUTINE TESTING

Routine tests are not made as frequently as they once were. It is generally agreed that the value of such tests in preventing or anticipating faults has been over estimated. Channel equivalents at a single frequency are checked weekly in most systems, but it is probable that this period will be increased to a month or more in the case of short circuits. System levels at important points, such as the output of the line amplifiers, carrier supply levels, synchronization, etc., are generally checked at intervals of 2 or 3 months. The emission currents of tubes are no longer measured, with the exception of the line amplifiers on multichannel systems, in which case the measurements serve as a check on intermodulation products. The frequency response of channels is at present measured at intervals of 4, or in some cases 12, months. Consideration is being given to the elimination of this test. Annual overhauls are no longer made, since a system with adequate maintenance should never deteriorate to the extent which would make such an overhaul necessary. If the performance of the system as a whole, as indicated by the measurements mentioned, is satisfactory, there is little to be gained by measuring its individual parts.

FAULT PROCEDURE

The efficient clearance of faults depends on both the technical ability of the personnel and the manner in which they co-operate with each other. To achieve the latter it is necessary that the procedure for receiving and passing on information relating to faults and their clearance be specified in some detail.

In South Africa there is a main or master control station which is advised of important breakdowns and which if necessary issues instructions in regard to repairs, authorizes work on important circuits, and collates and analyzes fault data. Several subsidiary control stations function in a similar capacity for their respective areas, usually in regard to secondary routes, and act as clearing centers for information required by the main control station. The fact that certain stations are designated as fault control stations does not lessen the responsibility of personnel, irrespective of the status of their station, to exercise their initiative, and to act immediately a fault occurs.

While the quantity of equipment in service has increased considerably, the supply of skilled maintenance personnel has not kept pace with this increase. Fortunately modern equipment requires very little maintenance. On the other hand, this means that the maintenance officer gets very little practice in clearing faults. Centralized repair depots and points from which technical assistance may be obtained telephonically are of considerable value.

ANALYSIS OF FAULTS

Control stations receive fault details daily, and in some cases weekly. The information is coded so that the clerical work involved is reduced to a minimum. The information is then transcribed to fault cards on a system and on a circuit basis. The control stations examine the fault returns and the fault cards at frequent intervals and any unusual trends are followed up. On a short-term basis this follow-up generally involves a high fault incidence at or on a particular station, system, or circuit. On a long-term basis the analysis reveals weaknesses in particular types of equipment.

An analysis of the outage of the 12-channel network over a period of a year is given in Table V. The line faults predominate and the following shows the number of faults per 100 miles of open-wire circuit per annum, based on 96,500 miles of circuit including a considerable amount on secondary routes and some on old routes:

Broken wires and binders	1.61
Protector faults	1.47
Faults due to any external objects except trees	0.96
Tree faults	0.53
Slack wires	0.38
Broken insulators	0.38
Wires cut by lightning	0.35
Faults due to workmen	0.14
Faults due to power contacts or surges	0.03
All other faults	0.96
Total	6.81

The contribution of line faults has been reduced very considerably by the provision of stand-by or reserve physical circuits for 12-channel systems. Power plant failures are relatively few in number but their contribution is high since one fault often affects a large number of circuits. A considerable number of the short breaks which occur appear to be due to unauthorized interference by maintenance staff, and it is important that stringent measures to prevent this be applied.

Protection from Lightning Damage and Interference

A considerable amount of lightning is experienced in South Africa: in several parts there are over 70 thunderstorm days per year.[18]

On a long open-wire line, such as a toll line, the crest surge current which occurs as a result of a near-by ground stroke rarely exceeds 200 amperes. The high currents associated with a high ground resistivity do not arise as they do on short customer lines. The current through the terminal protectors is in general of an oscillatory nature, partly because of successive reflections from the ends of the line which may increase the duration of the surge to about 1 to 5 milliseconds.

Air-gap metal-electrode protectors sealed in polyvinyl-chloride give very satisfactory service on toll lines. Gas protectors are not used in lightning areas because when damaged by lightning they give no indication of failure.

Toll protection at the end of an open-wire pair thus consists of two metal-electrode protectors with a drainage or equalizing coil. This latter coil is very effective in reducing the transverse voltage resulting from protector operation. No heat-coils or fuses are employed. It is important to bond the protector earth to the earth of the equipment, e.g., line filters, and also to the sheath of the lead-in cable. In high resistivity soil, where it may be difficult to obtain a low resistance earth connection, a horizontal buried wire provides the most economical solution.[19] In spite of the high lightning incidence and considerable amount of equipment installed in South Africa, the damage which occurs is negligible.

Apart from the effects of lightning discharges, protectors may operate due to changes in the electric gradient near the ground during periods of thunderstorm activity. In these cases the currents which pass through the protectors are less than a few milliamperes. The protector operation does, however, cause noise and this effect may be overcome by earthing the centerpoint of the line transformer, or by employing a drainage element such as silicon-carbide.

The effects of lightning on carrier telephone and voice frequency telegraph systems have been studied in some detail.[18] Briefly, the nature of the transients at the various points are determined by the filter characteristics. The frequency of oscillation is at the mid-band of a band-pass filter and at the cutoff frequency of a low-pass filter. The duration of the transient is a function of the bandwidth: the narrower the bandwidth the longer the duration. There is no appreciable lengthening of the transient until the output of the low-pass filter following the channel demodulator is reached. The important lengthening process occurs in the receiving telegraph filter, where the duration is always about 20 milliseconds for a 120-cycle spaced telegraph system. The peak voltage of the transient tends to increase as the gain of the terminal is increased.

It is found that amplitude-modulated voice frequency teletype circuits can be operated over open-wire carrier telephone channels with a negligible number of mutilations even under severe thunderstorm conditions, provided drainage (equalizing) coils are used, and provided that the system line levels are sufficiently high. The gain-time characteristics of the telegraph detector are also of importance. Operation over 12-channel systems is superior to operation over 3-channel systems.

Frequency-modulated telegraph systems are superior to amplitude-modulated systems as far as lightning mutilations are concerned but until their cost can be made comparable to that of amplitude-modulated systems, their use on high-quality carrier telephone channels is not justified.

References

1. Coaxial Cable Carrier Telephone Systems in South Africa, T. W. Elliott, N. J. Paola. *Transactions*, South African Institute of Electrical Engineers, Johannesburg, South Africa, vol. 43, Nov. 1952.

2. The Progress of South African Telecommunications, J. A. F. Michell. *Ibid.*, vol. 41,, Jan. 1950.

3. Trunk Telephone Practices in the Union of South Africa, P. Machanik. *Ibid.*, vol. 28, Mar. 1937.

4. The Evolution of the Long Distance Trunk System in South Africa, M. Hewitson. *Ibid.*, vol. 43, June 1952.

5. The Cape Town—Johannesburg Carrier Telephone and Telegraph System, P. Machanik, E. H. Harwood. *Electrical Communications*, New York, N. Y., July 1936.

6. Communication Network of the Union of South Africa, D. P. J. Retief. *Transactions*, South African Institute of Electrical Engineers, Johannesburg, South Africa, vol. 38, Mar. 1947.

7. SOJ-12 Open-Wire Carrier Systems in South Africa, D. P. J. Retief, H. J. Barker. *Electrical Communications*, New York, N. Y., Sept. 1947.

8. Twelve Channel Carrier Telephone Systems in South Africa, N. J. Paola, C. F. Boyce, I. C. Ramsay. *Transactions*, South African Institute of Electrical Engineers, Johannesburg, South Africa, vol. 39, Oct. 1948.

9. Automatic Trunk Exchanges and Two-Voice-Frequency Signalling in South Africa, D. W. Corlett. *Ibid.*, vol. 42, Nov. 1951.

10. Trends in the Design of Line Transmission Equipment, R. J. Halsey. *Paper 200*, Institution of Post Office Electrical Engineers, London, England, Nov. 1949.

11. Improved Equipment Practice Reduces Size of Telephone Transmission Systems, F. Fairley, R. J. M. Andrews, A. C. Delamare. *Electrical Communications*, New York, N. Y., Mar. 1950.

12. The Activities of the Power and Communication Systems Co-Ordinating Committee During the Period May 1941-July 1952, C. F. Boyce. *Transactions*, South African Institute of Electrical Engineers, Johannesburg, South Africa, vol. 44, July 1953.

13. Crosstalk in South African Telephone Lines, C. F. Boyce, N. J. Paola. *Ibid.*, vol. 33, Sept. 1942.

14. Line Development Problems in the Development of the 12-Channel Open-Wire Carrier System, L. M. Ilgenfritz, R. N. Hunter, A. L. Whitman. *Bell System Technical Journal*, New York, N. Y., Apr. 1939.

15. A 12-Channel Carrier Telephone System, B. W. Kendall, H. A. Affel. *Ibid.*, Jan. 1939.

16. Some Applications of the Type J Carrier Systems, L. C. Starbird, J. D. Mathis. *Ibid.*, Apr. 1939.

17. Maintenance of Multi-Channel Carrier Telephone Systems, F. O. Morrell. *Paper 194*, Institution of Post Office Electrical Engineers, London, England, April 1948.

18. The Protection of Open-Wire Communication Systems From Lightning Damage and Interference with Particular Reference to South Africa, C. F. Boyce, I. C. Ramsay, D. P. J. Retief. *Monograph*, South African Institute of Electrical Engineers, Johannesburg, South Africa, Apr. 1955.

19. The Earthing of Telephone Systems with Particular Reference to South Africa, C. F. Boyce. *Transactions*, South African Institute of Electrical Engineers, Johannesburg, South Africa, vol. 43, Dec. 1952.

Regulator Curves and Transient Currents of Double-Way and Double-Wye Rectifiers

L. E. JENSEN
ASSOCIATE MEMBER AIEE

C. E. RETTIG
ASSOCIATE MEMBER AIEE

CALCULATION of instantaneous and sustained current resulting from d-c faults on a power rectifier is extremely cumbersome. Certain idealized cases have been treated analytically, an important factor in establishing and understanding the circuit action. The analytical treatment also provides a necessary check on results obtained by other methods.

One effective way to reduce the amount of cumbersome calculation in such problems is to set up a model or miniature of the full-scale equipment which duplicates known conditions. This model or miniature can then be used to extend investigations through a great variety of circuit conditions with a considerable saving of time and effort over purely analytical treatment.

An extensive study of typical circuit conditions has been made with the transient network analyzer.[1] This study has resulted in a fairly complete coverage of regulation curves and current transients during d-c faults for most power rectifier operating conditions; and a simplified method of obtaining regulation and transient currents for typical power rectifiers.

Rectifier Circuits

The 6-phase double-way and 6-phase double-wye circuits used in the miniature study are shown in Fig. 1. The equivalence between double-way and double-wye rectifiers has been developed analytically in previous papers[2,3] and was verified by the use of the transient analyzer. Most of the data in this paper were taken with the double-way circuit.

Transient Analyzer Components

Germanium cells with the characteristics shown in Fig. 2 were used as the rectifying devices. Voltage and impedance magnitudes were chosen to make the voltage drop of the germanium cells negligible. The "arc-drop" of the cells was less than 0.3 per cent (%) of E_{d0} for the double-way circuit. The reactance elements had resistance-reactance (R/X) ratios in the order of 0.025. Thus,

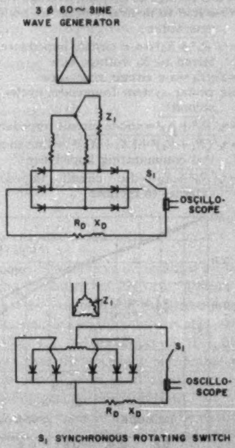


Fig. 1. Double-way and double-wye rectifier circuits

Paper 55-45, recommended by the AIEE Electronic Power Converters Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing November 29, 1954.

L. E. Jensen and C. E. Rettig are with the General Electric Company, Schenectady, N. Y.

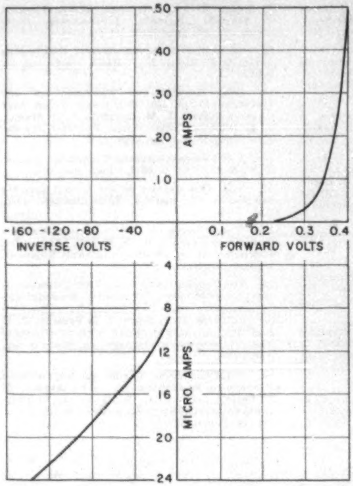


Fig. 2 (left). Germanium cell characteristics

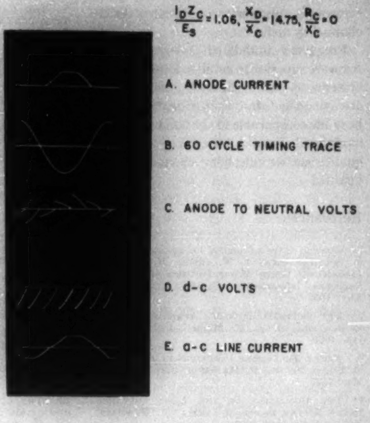


Fig. 4 (right). Change-over between second and third modes, inductive load

idealized conditions could be very closely approximated in checking against analytical work.

List of Symbols

I_d=steady-state d-c current, average amperes
i_d=instantaneous value of d-c current
i_1, i_2, etc.=instantaneous value of current in anode 1, anode 2, etc.
i_a, i_b, i_c=instantaneous a-c line currents
E_d=d-c terminal voltage, average value
E_{d0}=d-c voltage at no-load, average value
E_s=a-c line to neutral voltage at no-load, rms value
$Z_1=\sqrt{R_1^2+X_1^2}$=a-c circuit impedance referred to E_s voltage base
$X_1=2\pi fL_1$=a-c circuit reactance
f=a-c power system frequency, cycles per second
$Z_2=\sqrt{R_2^2+X_2^2}$=anode circuit impedance
$Z_c=\sqrt{(R_1+R_2)^2+(X_1+X_2)^2}$ = line-to-neutral commutating impedance
$Z_d=\sqrt{R_d^2+X_d^2}$–d-c circuit impedance external to rectifier

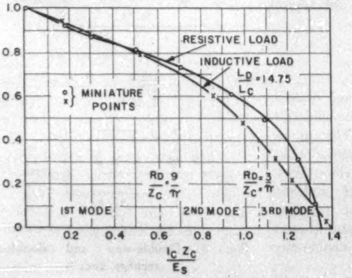


Fig. 3 (left). Comparison of miniature and theoretical regulation curve with commutating inductance only

Extended Regulation Curves

Regulation curves have been calculated by previous authors for the cases of highly inductive loads and of resistive loads, the commutating impedance in each case being purely reactive.[2–4] These curves are replotted in Fig. 3 along with points obtained from the miniature study showing the very close correspondence between theoretical and miniature curves. Oscillograms of voltages and currents, Figs. 4, 5 and 6, were taken of the points of change in mode of operation. These oscillograms also show excellent agreement with the theoretical wave forms.

Since rectifier resistance is omitted in these two idealized cases, the results are pessimistic. The curves indicate higher values of fault current than would be obtained in an actual rectifier. For example, calculations of regulation at rated load include a resistance loss component E_r as well as a reactance component E_x. The value of E_r is practically never less than one-third of E_x and in many cases it may be as high as two-thirds of E_x.

Figs. 7 through 10 show regulation curves for various values of R_c/X_c in comparison with the idealized curves. These curves were taken with both inductive and resistive loads. It can be seen that commutating resistance R_c reduces the difference between the cases of highly inductive and resistive loads. Also, the effect of R_c is of considerable importance. The range of the R_c/X_c ratio for practically all rectifier applications is between 0.11 and 0.43, shown in Table I of reference 5. This range is covered in Figs. 7 through 10.

For maximum usefulness it is desirable to have the regulation curve represented as simply as possible commensurate with desired accuracy. Fig. 11 shows straight-line approximations for the conditions covered in Figs. 7 through 10. The accuracy of this representation is within +9.0% or −2.5%. That is, a calculated value may be 2.5% lower than the actual current when E_d/E_{d0} is 0.3 and the R_c/X_c ratio is 0.12 with resistive load. The calculated value may be 9.0% higher than the actual current when $E_d/E_{d0}=0.55$ and the R_c/X_c ratio is 0.12 with inductive load. As the R_c/X_c ratio increases, the representation becomes more accurate.

Examination of arc-drop oscillograms taken under fault conditions indicate that the arc-drop is most accurately represented as a battery voltage. Since this voltage has no effect upon commutation, it may be considered as a single

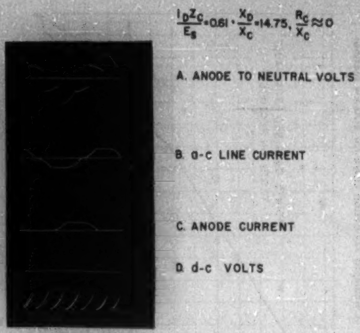


Fig. 5 (left). Change-over between first and second modes, inductive load

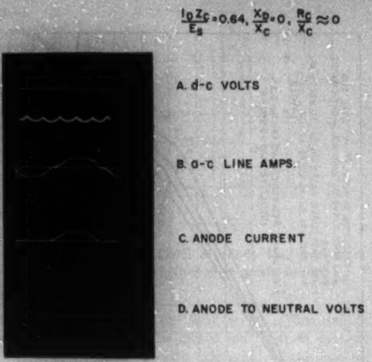


Fig. 6 (right). Change-over between first and second modes, resistive load

battery in the d-c bus. Typical values are indicated in Fig. 11. More precise values for particular applications depend upon the tubes or tanks used and the design value of short-circuit current. After obtaining a theoretical short-circuit current value from Fig. 11 a value of arc-drop may be selected from Figs. 8 or 9 of reference 6. This arc drop for the theoretical bolted fault conditions determines the maximum value of the straight-line approximation.

An approximate regulation curve for a particular application may be derived from Fig. 11. The intersection of arc-drop curve and appropriate regulation curve determines the value of bolted short-circuit current. The voltage at the break in the approximate regulation curve is reduced by the arc-drop at that current and the open-circuit voltage is reduced by the no-load arc-drop.

An example may clarify the procedure. Circuit constants of a 200-kw 300-volt sealed-ignitron rectifier, 6-phase double-wye connection will be used. Actual fault tests made on this rectifier will then serve as a check on the accuracy of the method. The circuit constants are as follows (commutating impedance shown in Table I).

Table I. Commutating Impedance

	X_c, Ohms	R_c, Ohms
Supply lines	0.026	0.0032
Rectifier transformer	0.060	0.0154
Anode leads	0.003	0.0030
	0.089	0.0216

$$Z_c=\sqrt{R_c^2+X_c^2}=0.0916 \text{ ohm per phase}$$

$$\frac{R_c}{X_c}=0.243$$

Transformer voltage is

$E_s=297$ volts line-to-neutral

Theoretical no-load voltage is

$E_{d0}=1.17\ E_s=348$ volts

Theoretical bolted fault current per wye is

$$I_c=1.41\times\frac{E_s}{Z_c}=1.41\times\frac{297}{0.0916}=4{,}570 \text{ amperes}$$

Arc drop is, at no load

$$E_a=15 \text{ volts}, \frac{E_a}{E_{d0}}=\frac{15}{348}=0.043$$

and at 4,570 amperes

$$E_a=20 \text{ volts } \frac{E_a}{E_{d0}}=\frac{20}{348}=0.057$$

D-c impedance is shown in Table II.

Table II. D-C Impedance

	X_d, Ohms at 60 Cycles	R_d, Ohms
Interphase transformer	0.0018	0.0008
Lines to short-circuiting breaker	0.0300	0.0025
	0.0318	0.0033

The effect of the interphase transformer resistance is small but may be accounted for by adding its per-unit voltage drop at the theoretical bolted fault current to the arc drop. Thus

$$\frac{E_a'}{E_{d0}}=\frac{E_a}{E_{d0}}+\frac{2\times1.41E_s}{Z_c}\times\frac{R_d}{E_{d0}}=0.057+$$
$$9{,}140\times\frac{0.0008}{348}=0.057+0.021=0.078$$

The effect of interphase transformer inductance and bus inductance with this R_c/X_c ratio is small, and will be neglected in this example.

The modified arc-drop line may be constructed on Fig. 11 by joining the two points $(I_cZ_c/E_s=0)$, $(E_a/E_{d0}=0.043)$ and $(I_cZ_c/E_s=1.41)$, $(E_a/E_{d0}=0.078)$. The intersection of this line and the regulation curve for $R_c/X_c=0.243$ (interpolated midway between the curves for $R_c/X_c=0.21$ and $R_c/X_c=0.30$) determines the bolted fault current. In this case $I_cZ_c/E_s=1.33$. Therefore the sustained bolted fault current is

$$I_d=2\times1.33\times\frac{E_s}{Z_c}=2\times1.33\times\frac{297}{0.0916}=8{,}620 \text{ amperes}$$

The no-load voltage is

$E_{d0}-E_a=348-15=333$ volts

The point at which the regulation curve breaks, located by the co-ordinates, is

$$\frac{I_cZ_c}{E_s}=0.93$$

$$\frac{E_d}{E_{d0}}=0.48-0.07=0.41$$

or

$$I_d=2\times0.93\times\frac{E_s}{Z_c}=2\times0.93\times3{,}240=6{,}030 \text{ amperes}$$

$E_d=0.41\times348=143$ volts

The regulation curve determined by these points is shown in Fig. 12. Points obtained by oscillograph measurements of faults on this equipment are also shown in Fig. 12. Observation of this figure shows that the accuracy obtainable with these curves is as good as the precision in oscillographic measurement.

In some cases it may also be desirable to include the effects of anode circuit inductance. Regulation curves including the effect of anode circuit inductance are

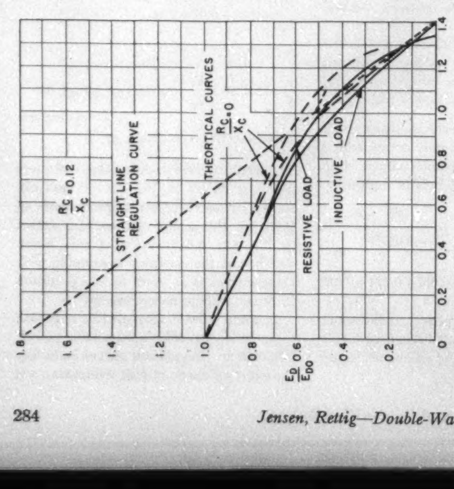


Fig. 7 (left). Regulation curve including commutating resistance $R_c/X_c=0.12$

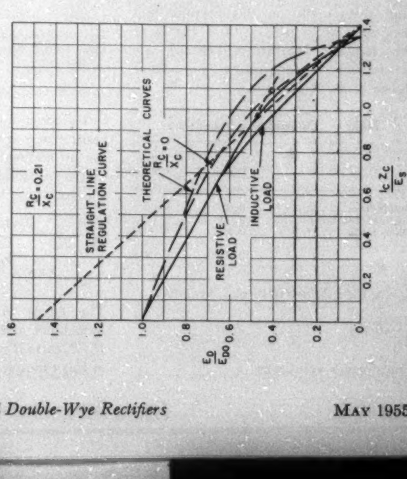


Fig. 8 (lower left). Regulation curve including commutating resistance $R_c/X_c=0.21$

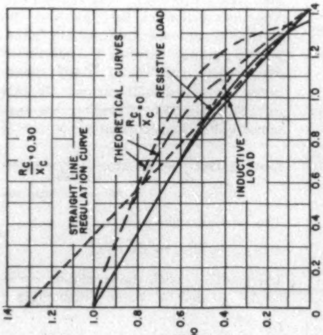


Fig. 9 (above). Regulation curve including commutating resistance $R_c/X_c=0.30$

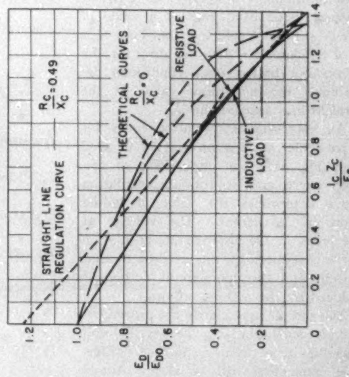


Fig. 10 (below left). Regulation curve including commutating resistance $R_c/X_c=0.49$

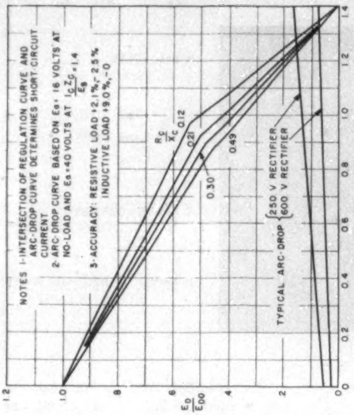


Fig. 11 (above). Straight-line regulation curves

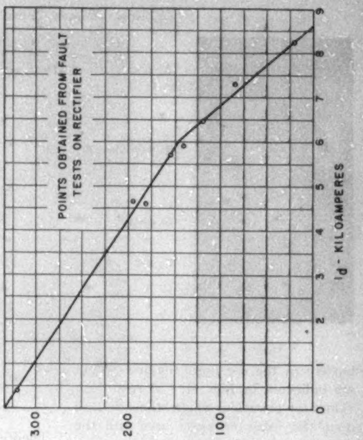


Fig. 12 (below). Calculated regulation curve for 200-kw rectifier

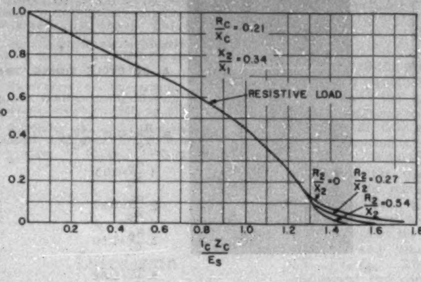


Fig. 13 (left). Effect of anode lead impedance

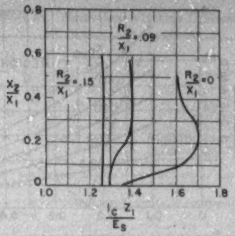


Fig. 14. Effect of anode lead impedance on bolted short-circuit current

shown in Fig. 13 for the case of $X_2/X_1=$ 0.34. The inductance added in series with each rectifying element could represent secondary leakage plus lead inductance. For the majority of the regulation curves the reactance X_2 adds directly to X_1. However, for the bolted short-circuit condition X_2 has the effect of increasing the steady-state fault current, as is shown in Figs. 13 and 14.[6]

The explanation of this phenomenon is given in reference 6. Inductance in series with each anode extends the time of conduction of each anode beyond 180 degrees at the bolted-fault condition, thereby resulting in higher d-c short-circuit current. Practical values of R_2/X_2 are not usually lower than 0.22. The combined effect of arc-drop and R_2/X_2 ratio is such that the actual current will not exceed the straight-line approximation.

Transient Fault Currents

When a short circuit occurs on a d-c circuit of a rectifier, a peak current greater than the steady-state fault current may or may not occur, depending on the relative values of the circuit impedances. The miniature circuits of Fig. 1 were also used to study transient behavior of the short-circuit current as a function of circuit impedances. Switch S_1 of Fig. 1 is a rotating switch driven by a synchronous motor. Another switch driven by the same synchronous motor was used to trigger the sweep circuit of the oscilloscope. Thus, switch S_1 could be closed and opened, in synchronism with triggering of the oscilloscope, several times a second so that the transient current trace appeared to stand still on the oscilloscope screen. Pictures of the traces were taken with a special camera when desired.

The wave shape of the transient currents are shown in Fig. 15. The peak current is limited by either resistance or reactance in the fault, or by both, as shown in Fig. 16.

For the case of a bolted fault or a resistance fault the current during the transient interval is made up of two components. One is the sustained current, the other is an a-c component of power system frequency which decays in

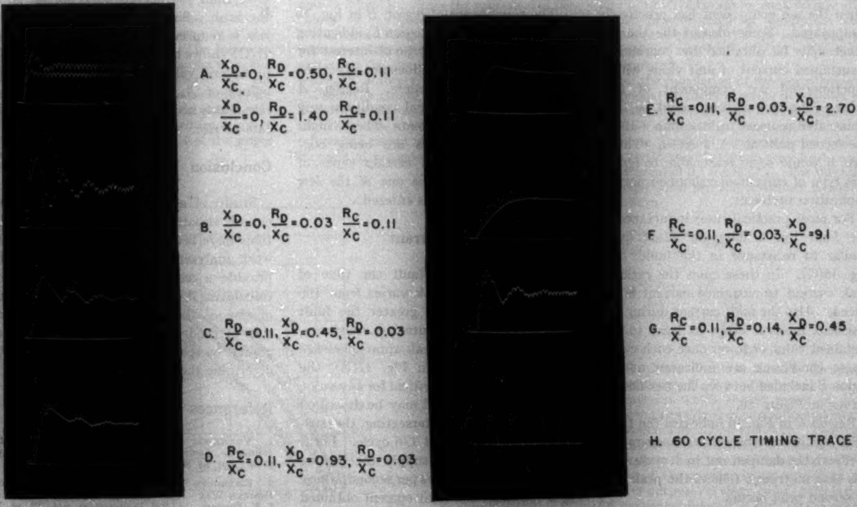


Fig. 15. Oscillograms of typical transient currents with impedance in fault

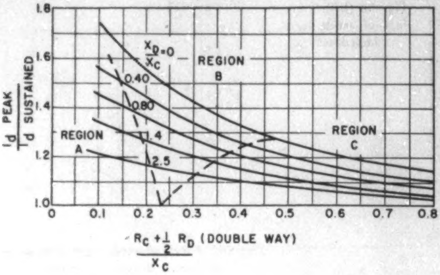


Fig. 16 (above). Effect of impedances on peak current

an exponential manner. The larger the resistance to reactance ratio, the more rapidly the exponential decays. See Fig. 15(A) and (B). Since the a-c component must be equal to the sustained component at the time the fault is applied, the first crest of current occurs approximately 1/2 cycle after the fault application. Considering the extreme condition where $R_c/X_c=0.12$ with bolted fault, the peak current is 1.68 times the sustained current. Except for the stresses developed in the rectifier transformer, the principal interest in this peak current is in the thermal effects. The influence on d-c circuit breakers would be thermal. The breakers do not interrupt until the ensuing trough of current is reached at the end of 1 cycle (for high speed breakers) or until 2 or more cycles have elapsed (for medium- and low-speed breakers) when the a-c component has practically disappeared. Some idea of the thermal effect may be obtained by considering a sustained current of unit value with a superimposed a-c component of 0.68 value. The rms current would be 1.21. This value reduces to less than 1.10 on the second peak at 1 1/2 cycles. Therefore, it would seem reasonable to ignore this type of current overshoot in normal application problems.

For most practical cases inductance in the fault has a damping effect quite similar to resistance in the fault. See Fig. 15(G). In these cases the ratio of peak current to sustained current is reduced. Also the fault current during the transient interval drops at least to the sustained value or lower once each cycle. These conditions are indicated as the region *B* included between the two dotted curves in Figure 16.

Region *C* in Fig. 16 indicates the transient conditions where a peak occurs but is effectively damped out in 1 cycle. In this case no trough follows the peak and no second peak occurs.

Region *A* in Fig. 16 indicates the transient conditions where the d-c circuit has a long enough time constant that the d-c fault current does not follow the a-c currents into the trough. The dividing line was chosen for the condition wherein the current troughs following each peak were equal to or greater than the sustained value.

The significance of region *B* in Fig. 16 has been discussed. Region *C*, indicating a 1/2-cycle peak, may be of interest for the thermal effect but does not affect the interrupting requirements. Region *A* may affect both thermal conditions and interrupting requirements where high-speed circuit breakers are being considered. Flashover of certain types of d-c machines would be one of the few cases where region *A* is entered.

Rate of Rise of Current

With a resistance fault the time of occurrence of the peak varies from 180 to 120 degrees. The greater the fault resistance and commutating resistance, the more nearly the peak approaches 120 degrees. As shown in Fig. 17(B), the initial rate of rise is identical for any value of fault resistance, and may be described as a straight line intersecting the sustained fault current at 1/6 cycle. For a 60-cycles per second supply the rate of rise is 360 × I_d amperes per second, where I_d is the sustained fault current obtained from Fig. 11.

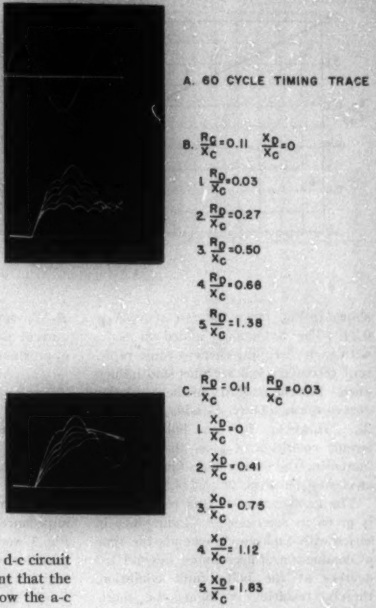


Fig. 17 (right). Oscillograms of transient currents with peaks with inductance in fault

With both resistance and inductance in the fault, the peak current may occur later than 180 degrees after initiation of the fault. Similarily, the initial rate of rise is reduced by the inductance. Fig. 17(C) shows the magnitude of this effect for several values of X_d/X_c. For larger values of X_d/X_c the ratio E_{d0}/L_d is a reasonably accurate measure of the initial rate of rise.

Conclusion

Study of a great number of rectifier circuit variations under d-c fault conditions have been made on a transient network analyzer. The results of this study provide a reasonably simple method of calculating d-c fault currents and the rate of rise of these currents. Accuracy of this method is commensurate with the ability to specify the impedances which determine these currents.

References

1. TRANSIENT ANALYZER APPLICATIONS, I. B. Johnson. *General Electric Review*, Schenectady, N. Y., vol. 54, Sept. 1951, pp. 22–29.

2. EXTENDED REGULATION CURVES FOR 6-PHASE DOUBLE-WAY AND DOUBLE-WYE RECTIFIERS, I. K. Dortort. *AIEE Transactions*, vol. 72, pt. I, May 1953, pp. 192–202.

3. INFLUENCE OF A-C REACTANCE ON VOLTAGE REGULATION OF 6-PHASE RECTIFIERS, R. L. Witzke, J. V. Kresser, J. K. Dillard. *Ibid.*, pp. 244–53.

4. EQUIVALENT MACHINE CONSTANTS FOR RECTIFIERS, I. K. Dortort. *AIEE Transactions*, vol. 72, pt. I, Sept. 1953, pp. 435–38.

5. PROTECTION OF ELECTRONIC POWER CONVERTERS, AIEE Committee Report. *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 813–29.

6. RECTIFIER FAULT CURRENTS—II, C. C. Herskind, A. Schmidt, Jr., C. E. Rettig. *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 243–52.

◆

Discussion

I. K. Dortort (I-T-E Circuit Breaker Company, Philadelphia, Pa.): This paper provides another excellent experimental check on the theoretical work of previous authors. In addition, it offers a practical means of attacking the problem of real rectifiers under transient conditions. Using the data presented, I have rechecked the approximate method which is given in Fig. 1 of reference 4 of the paper for the case of noninductive loads.

The derived regulation curves given in Fig. 1 of that paper are essentially correct except for an error in the resistance line. In equation 7 of that paper it was assumed that the copper losses of the rectifier transformer and anode cables would increase as the square of the direct current I_d. Because of the change of wave shape of the anode currents, which become sinusoidal as we approach short circuit, the copper losses increase at a lesser rate and the equivalent voltage drop increases less rapidly than the direct current. In terms of the load-loss P_r caused by anode currents of normal wave shape flowing through the total commutating resistance R_c at rated direct current I_{dn}, sinusoidal anode currents adding up to the same direct current would result in a load loss of $\pi^2 P_r/12 = 0.822\ P_r$. The losses in the d-c circuits, interphase-transformer, and arc, are not materially affected by the anode current wave shape. Assuming, for the sake of simplicity, that the anode currents are sinusoidal over the entire range of load, equation 7 of reference 4 of the paper becomes

$$E_r = E_{ao} + mR_a I_d + \frac{0.822 P_{r(ac)} I_d}{I_{dn}^2} + \frac{P_{r(dc)} I_d}{I_{dn}^2}$$

In Fig. 18 the first curve represents the theoretical regulation curve of an idealized rectifier and the straight resistance lines (approximations) are based on R_c/X_c ratios of 0.12, 0.21, 0.30 and 0.49 as used in the four corresponding tests of the paper under discussion. Subtracting these resistance lines from the theoretical no-loss curve results in the regulation curves for these four conditions. Superimposed on them are heavy curves transcribed from the test curves given by the authors in Figs. 7, 8, 9, and 10, and transformed to the base $(X_{cpu})(I_d/I_{dn})$. The areas between the theoretical and corresponding test curves have been cross-hatched and it can be seen that the difference is never greater than $2^1/_2\%$ in the vicinity of short circuit.

The agreement for cases $R_c/X_c = 0.12$ and 0.21 is almost perfect. The larger discrepancies in the curves for $R_c/X_c = 0.30$ and 0.49 have not been accounted for. Considering the simplifying assumptions used to derive the theoretical curves and the poor resolution of the published test curves used for comparison, the over-all verification is almost startling.

The results of the tests made by the authors on the 200-kw test rectifier were handled the same way with excellent results. The resistance line and derived regulation curve are shown dotted on the accompanying illustration.

The points depicted by small circles are taken from the test points of Fig. 12 of the paper. It will be noted that the value 1.0 for S gives a bolted short-circuit current of 8,200 amperes, whereas the approximate method used by the authors comes to 8,620 amperes. In Fig. 12 the last test point shown occurs at 8,200 amperes and approximately 20 volts. Since the resistance of the cables to the short-circuiting breaker is 0.0025 ohm it seems that the voltage readings were taken at the rectifier terminals and that the actual short-circuit current was 8,200 amperes.

From the circuit constants given by the authors for the 200-kw test set, it appears that this is the same equipment described in reference 5 except for minor changes and the increase of secondary voltage from 283 to 297 volts. The actual short-circuit current obtained in the earlier tests was 7,700 amperes. Correcting the 8,200 amperes predicted by the theoretical curve by the ratio of secondary voltages, the method I have proposed yields a bolted short-circuit current of 7,800 amperes.

Considering the inaccuracies contained in the curves of Fig. 11, it seems a pity the authors did not incorporate average arc-drop characteristics in the curves, instead of requiring a separate operation to obtain equivalent constants. Applying the two arc-drop lines for 250- and 600-volt rectifiers to Fig. 18, the maximum difference in bolted short-circuit current is only $4^1/_2\%$ when tested on the least favorable curve $R_c/X_c = 0.49$.

The deviation between rectifiers in any one class would of course be much smaller, and a family of curves for each voltage class would provide accurate equivalent constants at a glance. To the same end, use of a rational interval for R_c/X_c, will facilitate interpolation.

For the calculation of the contribution of individual rectifiers to faults in complex d-c systems, a table of equivalent machine constants may be even more useful. As an example, Table III was constructed for 300-volt rectifiers with noninductive loads, using an interval of 0.1 for R_c/X_c. For this table the equivalent per-unit resistance is derived as follows

$$\mathcal{R} = \frac{\Delta E_d}{\Delta I_d} = \frac{\Delta E_{dpu}}{\Delta S} X_{cpu} \frac{E_{do}}{I_{dn}}$$

$$\mathcal{R}_{pu} = \frac{\Delta E_{dpu}}{\Delta S} = \frac{\mathcal{R}}{X_{cpu}\dfrac{E_{do}}{I_{dn}}}$$

Table III. Equivalent Machine Constants

R_c/X_c	$\mathcal{E}_{1pu}$	$\mathcal{R}_{1pu}$	S_1	$\mathcal{E}_{2pu}$	$\mathcal{R}_{2pu}$	S_2
0	0.95	0.50	0.82	2.16	1.97	1.10
0.1	0.95	0.58	0.80	1.89	1.74	1.08
0.2	0.95	0.67	0.79	1.67	1.57	1.05
0.3	0.95	0.75	0.77	1.49	1.45	1.02
0.4	0.95	0.84	0.74	1.35	1.38	0.98
0.5	0.95	0.93	0.70	1.25	1.35	0.93

$\mathcal{E}_{1pu}$ = equivalent machine per-unit voltage for first phase

$\mathcal{R}_{1pu}$ = equivalent internal per-unit resistance for first phase

$S_1 = X_{cpu}\dfrac{I_d}{I_{dn}}$ at intersection of first and second phase

$\mathcal{E}_{2pu}$ = equivalent machine per-unit voltage for second phase (near short-circuit)

$\mathcal{R}_{2pu}$ = equivalent internal per-unit resistance near short circuit

$S_2 = X_{cpu}\dfrac{I_d}{I_{dn}}$ at short-circuit

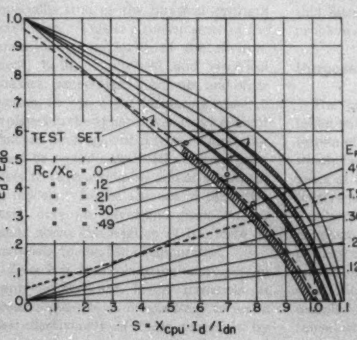


Fig. 18. Comparison of theoretical and test curves

W. R. Crites (General Electric Company, Schenectady, N. Y.): This paper presents information which should prove of great value in the analysis of current transients on d-c systems involving rectifiers. The regulation curves permit the determination of an equivalent resistance which can be used to represent the rectifier in an equivalent circuit for purposes of short-circuit calculation. Of particular interest is Fig. 16 concerning the possible occurrence of a peak on the d-c fault current and the magnitude of that peak.

A knowledge of the peak current is important in determining the adequacy of circuit breakers, busses and other system components. In this respect, the authors have stated that high-, medium-, and low-speed circuit breakers will not interrupt the current until after the peak value has been passed. This is true of the medium and slow-speed circuit breakers; however, there are high-speed breakers available which may operate in approximately one-half cycle and which would in that event be required to interrupt the peak current. The slower circuit breakers must pass the peak current and are therefore subjected to the forces associated with it even though they may not be required to interrupt it. Busses and other circuit elements through which the current passes will also be subjected to the forces of the peak current. The effect of these forces on the system components depends on the physical construction of the components, that is, such factors as mechanical strength and resonant frequencies.

In consideration of these factors, it appears that the authors' conclusion, "Therefore, it would seem reasonable to ignore this type of current overshoot in normal application problems," requires fuller explanation than is given in the paper.

L. E. Jensen and C. E. Rettig: The excellent agreement with the analysis by Mr. Dortort helps to show the effectiveness and accuracy of the miniature circuit used in this study, thus demonstrating the usefulness of the transient analyzer.

The approximate regulation curves of Fig. 11 were intended to show how assumptions could be made to simplify the calculation of short-circuit currents. Arc-drop effects could be included in the regulation curves so that fault currents could be read directly but it was felt that it would be better to keep the arc drop separate so that the reader could see at a glance what approximations were being made. Further simplification may be desired in some cases while in others greater accuracy may be desired. All of the curves for the various conditions have been included in the paper so that the reader may obtain the degree of precision which is desired for his particular needs.

The authors are aware that some high-speed breakers may interrupt near the time of occurrence of the transient peak of current. It was not the intention to imply that this overshoot should always be ignored. Later in the paper, information was included so that the reader could reach the degree of accuracy desired. Thus the curves of Fig. 16 which were of particular interest to Mr. Crites were included so they could be used for just such conditions as he suggests.

Significant Factors in Thermal Aging Tests on Flexible Sheet Insulation

T. W. DAKIN
ASSOCIATE MEMBER AIEE

H. M. PHILOFSKY
ASSOCIATE MEMBER AIEE

W. C. DIVENS
STUDENT MEMBER AIEE

THE revision of the present temperature classification of insulating materials as in AIEE Standard No. *1*[1] seems increasingly necessary to accommodate insulation application practice and apparatus temperatures to the higher temperature at which some of the newer synthetic materials seem to be safely operable. It has been the goal of the Working Group on the Evaluation of the Thermal Stability of Insulating Materials which was set up in the Basic Science Subcommittee on Dielectrics to develop acceptable standard methods of evaluating the thermal stability of insulating materials. Such methods should put any revision of the present standards onto a rational and conservative basis.

An acceptable method of evaluation of the life of an insulating material should have a proven relation to the operation of this insulation in apparatus. An attempt should be made to avoid arbitrary criteria of life end points. It is desirable to have a clear-cut general rule, but this should not be at the sacrifice of correlation with practical experience and the required function of the insulation. For example, it is not logical to rule the insulation dead when it has lost 50 per cent of its initial dielectric strength if the application required only 5 per cent of it.

In many cases the engineers who design apparatus and apply insulation cannot give a clear-cut answer in numbers to the question of what properties are required of the insulation they use. This is especially true of the physical strength properties. This is a consequence of the tremendous complexity of the environment, both of physical and electrical stresses, temperature gradients, and other factors, in which the insulating material is placed in the apparatus. It is apparent that the development of tests on materials must be co-ordinated with at least some tests to destruction on actual apparatus or models. In the absence of the latter tests, only an intelligent guess can be made about the relation of life tests on materials to the life of apparatus.

Since it is desirable to make progress in applying new and better materials as soon as possible, it is necessary to prove the effectiveness of these materials by accelerated tests. The most reasonable way of accelerating the degradation of insulation to determine its life, and, at the same time, to obtain information about its expected life under unaccelerated conditions, is to raise the temperature. It had already been pointed out[2] that chemical reaction rate principles apply to insulation aging and that these principles are used to extrapolate the results of short-time tests at high temperatures to longer times at more normal temperatures. This extrapolation is made by using the relation between temperature and life at some arbitrary end point, i.e.

$$\log (\text{life}) = k_1 + k_2/T$$

where T is the absolute temperature, and k_1 and k_2 are constants characteristic of the material being tested and the test conditions. The accuracy and applicability of this relation has been shown in a good many tests. However, to sound a note of caution, there are some situations which are too complex for it to fit. It is best that it be checked in each new situation by determining the linearity of a plot of log (life) versus $1/T$ by measuring the life a a series of a least three different temperatures in accelerated tests.

It should be thoroughly understood, however, that an accurate extrapolation from an accelerated high-temperature test does not of itself give a correct and useful result unless the accelerated test accurately represents the aging-to-failure conditions in the application. Satisfying this latter condition is the crux of the problem.

Purpose of Present Investigation

The present investigation was undertaken with the object of evaluating a life test method for varnished glass cloth which had been in use for a number of years at the Westinghouse Research Laboratories. The test was very simple: a 4-mil standard weave glass cloth was impregnated by a motor-controlled dipping process with the varnish baked, given a second dip, and finally cured. The initial dielectric strength of a representative sample of the varnished sheet was obtained and aged in an oven at a constant high temperature while held flat between layers of glass cloth. It was removed periodically from the oven to be proof-tested at room temperature at 50 per cent of its original dielectric strength. A 2-inch electrode was applied momentarily in succession to a number of different spots on the sheet so that a sizeable fraction of the total area was tested each time. Several 6- by 6-inch sheets were usually tested in parallel. The end point was arbitrarily set when half the spots tested failed.

The 50-per-cent dielectric strength criterion of end point has been used extensively;[3-6] sheet or tape insulation directly applied to conductors were tested and, in most cases, subsequently dipped in varnish. Such tests are more nearly representative of the application, but they involve variable factors peculiar to the manner of wrapping or taping or varnish treating which affect the observed life but may not be characteristics of the original tape or sheet material if it were applied to the conductor differently. It is likely that both types of tests should be done.

A critical analysis of this life test method seemed in order, to determine the real value of such a test for the purpose of measuring the operating temperature limits of insulating materials. It was felt that this analysis would spotlight a number of points of general significance to the whole program of insulation material life testing.

Such an analysis can be begun by asking several salient questions: what was really being tested—the relative life of the varnish to some other varnish or the absolute life? or a much more restricted thing—only the life of this particular composite insulation to a particular level of dielectric strength when aged in a particular way? The possible correlation of this test with the expected life of similar insulation in apparatus, and the correlation of this with other property tests should be looked into. The failure mechanism and criterion of failure as well as a number of factors associated with the test procedure need to be examined for their effect on the result.

Factors Affecting Life Test Results

It was known almost from the initial use of the test that varnishes which were

---

Paper 55-37, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing November 17, 1954.

T. W. DAKIN, H. M. PHILOFSKY, and W. C. DIVENS are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

believed from general experience and other physical tests to be more thermally stable usually showed longer life on the test. For example, the silicone varnishes showed vastly longer life at comparable temperatures than the conventional organic varnishes. In fact, it proved to be impractical to test most organic varnishes at the same temperature as silicone varnishes, their apparent lives on the test being so vastly different. Thus, the value of the test in the laboratory seemed to be chiefly in its ability to give a reproducible and positive differentiation between different varnishes. This is illustrated in Fig. 1, where the lives of two different organic varnishes and two different silicone varnishes are graphed as a function of temperature.

These data in Fig. 1 indicate the differences in life of organic materials which, in the present classification system, would be rated exactly the same. It is indicated that insulation incorporating varnish *A* could be operated about 20 degrees higher than varnish *B* with a comparable life. Similarly, more than a 30 degree higher operating temperature would be possible with the better silicone varnish *A* over the silicone varnish *B*. Both of these are now restricted to 180-degree (Class H) temperatures. Tests such as these indicate that 180 degrees is too high a temperature for some poorer silicones but much too low for some of the superior silicones. Accelerated test data, shown in Fig. 1, can be used by extrapolation to predict the life at lower operating temperatures.

In discussing the variation of the life test results with different varnishes, the mechanism of failure should be considered. The test procedure measures the time for the dielectric strength to diminish to a certain level. The dielectric strength of varnished glass is largely caused by the varnish, the glass serving only as a skeleton spacer. The dielectric strength may decrease in the aging process by several means: loss in thickness of the varnish layer by evaporation (of decomposition products or initial low molecular weight constituents), by cracking of the varnish, or by a change in the chemical composition of the varnish which changes its dielectric strength for the same thickness. The last named effect might, on theoretical grounds, even be expected to increase the resin dielectric strength, unless the resin power factor is increased abnormally. Experience in examining the test specimens indicates that cracking is usually the most important factor in decreasing the dielectric strength, although evaporation can be quite an appreciable factor with some organic resins. Thus, the test method measures the time for an arbitrary extent of cracking to occur together with, in some cases, an amount of evaporation which collectively reduce the effective thickness of the insulation by half. The cracking results from a decrease in flexibility and shrinkage of the resin, probably caused by polymer cross-linking reactions. This makes the resin incapable of accommodating itself without rupture to the stresses imposed by its differential expansion and contraction in the glass cloth framework.

It should be recognized that the test itself sets several arbitrary conditions which may not be representative of the varnished glass itself, or the varnish separately in some other way. The life criterion is dielectric failure at a level (over 500 volts per mil) which is considerably higher than is normally imposed on insulation of this sort. There are some applications of varnished glass insulation where the glass serves only as a spacer and the varnish needs to add no dielectric strength. It is difficult, if not impossible, to see how the test under discussion is a measure of the varnish life where its function is in the spacer category. The manner in which the varnished glass samples are heated is also very arbitrary. In many applications at first, the varnished glass might be flexed considerably when put into place; during operation it might be flexed only slightly. It might be abraded to some extent or subjected to the possible corrosive effects of metals with which it comes into contact in a machine. All of these are factors which play no part in the test method, but which might be very important factors in apparatus.

Mathes and Morgan[7] and Plettner and Currin[8] have suggested that the varnish should be elongated by mild flexing, by about 1 per cent to simulate strains occurring in service. Experience has indicated that such flexing (over a 1-inch mandrel) reduces the observed varnish life on glass cloth to about half. The introduction of an arbitrary small amount of flexing in such a test can be rationally justified and should be more thoroughly explored to determine the amount most similar to the application conditions.

That the test measures the relative life of the combination rather than the varnish itself is shown in Fig. 2. Here the base materials (glass, cloth, mica paper, and asbestos paper, each 2 mils thick) are compared, each impregnated with the same silicone varnish to a total thickness of 5±1 mil. All of the tests to be discussed from this point were made with this thickness material. The difference in the dielectric strength life between the glass cloth and the other base materials is considerable.

The mica and asbestos papers also show a different type of trend of dielectric strength with aging time with a silicone

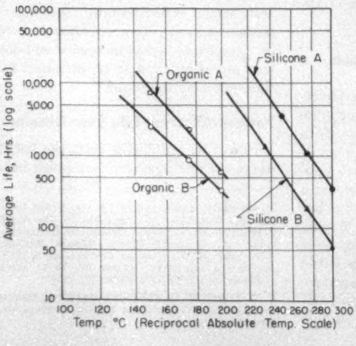


Fig. 1 (left). Thermal life of varnishes on glass cloth, 50-per-cent dielectric strength failure

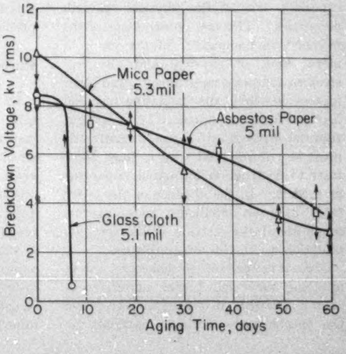


Fig. 2 (right). Effect of aging at 300 degrees centigrade on the dielectric strength of several base materials impregnated with a silicone resin

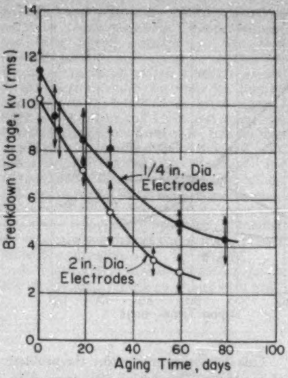


Fig. 3. Effect of aging at 300 degrees centigrade on the dielectric strength of mica paper impregnated with a silicone resin

varnish. The glass cloth exhibits, after a period of time, a sudden drop in dielectric strength from near the original value to a low value, in many spots about that of air, while the other base materials demonstrate a more gradual decline in dielectric strength. In the case of the glass cloth, cracking can occur at one time completely through the thickness of the varnish in streaks parallel to and between the glass threads.

This behavior of glass cloth with this silicone resin dispels any question of what level of dielectric strength should be regarded as a failure point. The conventional organic resins and some other silicone varnishes which have been tested with glass cloth do not show the sudden decline in dielectric strength but demonstrate a more gradual decline.

The mechanical stresses imposed on the varnish appear to be less in the mica paper and asbestos paper where the mineral particles are weakly held to each other and are able to expand and contract with the varnish. The paperlike structure seems to prevent cracks extending at one time completely through the sheet. The mica paper shows an aging characteristic peculiar to it: the varnish layer on either side of the mica paper cracks separately into small scales which tend to curl up. This does not seem to affect the integrity of the impregnated mica beneath.

Another important factor in the results obtained on a dielectric strength test of this sort is the electrode area. The electrode area is especially important when testing a nonuniform material having defects such as mentioned in the foregoing. The cracking phenomenon is a random effect which in itself introduces defects of varying depth. These will show a distribution in size with a few very deep cracks This leads to a more rapid decline in dielectric strength values with large electrodes, as shown in Fig. 3. The larger defects are more likely to be encompassed by a large electrode.

Application of data from tests of this sort to predict the behavior of similar insulation in an apparatus where much larger areas are involved must be done with care to consider the distribution of test results If a good distribution curve is obtained by a large number of tests with small electrodes, the probable breakdown value with an area equal to that exposed in a machine can be predicted. The distribution curves for the breakdown of aged and unaged silicone-impregnated mica paper are shown in Fig. 4. The distribution of breakdown voltages with 1/4-inch electrodes and with 2-inch electrodes are compared, with the voltages plotted on a logarithmic scale to emphasize the large percentage difference between the tests with 2-inch electrodes and 1/4-inch electrodes on the aged material as compared to the unaged. These breakdown voltage data do not show a normal distribution when plotted either on a linear or a logarithmic scale versus the normal probability integral scale used on the abscissa in Fig. 4. To follow a normal distribution, the data would need to fall on a straight line on such a plot.

High humidity is frequently a factor encountered with machines in service. It is important when following a shutdown period where the machine has been allowed to cool and the insulation absorbs moisture. In Fig. 5, the dielectric strength and dissipation factors of an oil-modified alkyd varnish on asbestos paper are shown as they vary with aging time and after exposure to high humidity. The effect of the moisture absorption after aging is quite appreciable on both the dissipation factors at 35- and 90-per-cent relative humidity respectively. In fact, the dissipation factor at 90-per-cent relative humidity rose so high as to become impractical to measure with accuracy. The dielectric strength at the high humidity was remarkably lower, probably as a result of dielectric heating of the sample during the test. These tests were made after 20 hours of humidification and the dielectric strength was determined with a 2-kv-per-second rate of voltage rise.

The rate of moisture absorption on exposure to high humidity is an important factor in considering the effect of such exposure. In many situations, equilibrium may not be reached for weeks. This is especially true of thick insulation. Even in the case of the 5-mil insulation

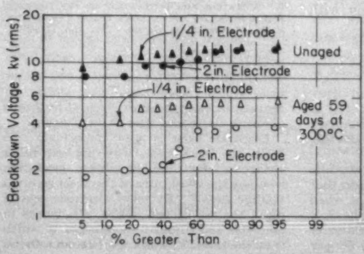
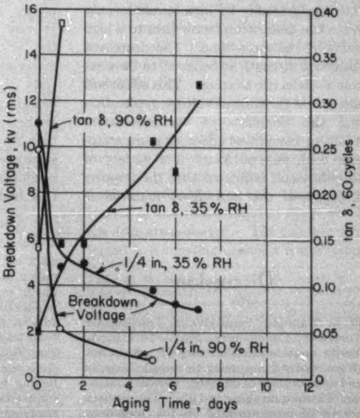


Fig. 4 (left). Distribution of breakdown voltage of silicone-impregnated mica paper

Fig. 5 (right). Effect of aging at 225 degrees centigrade on the electrical properties of an oil-modified alkyd varnish on asbestos paper



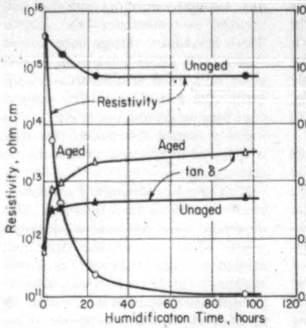
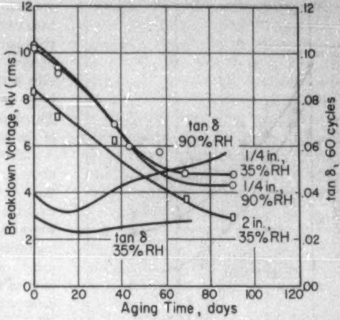


Fig. 6 (left). Effect of humidification on electrical properties of a silicone varnish on mica paper unaged and aged 30 days at 300 degrees centigrade

Fig. 7 (right). Effect of aging at 300 degrees centigrade on electrical properties of a silicone varnish on asbestos paper



tested here, the equilibrium is not quite complete in 24 hours. This is illustrated in Fig. 6, with data on silicone-varnish-impregnated mica paper.

Silicone-varnish-impregnated insulation of this sort appears to be much less moisture-sensitive, even at equivalent dielectric strength conditions, than conventional organic-varnish-impregnated insulation. This is illustrated by a comparison of Fig. 5 with Figs. 6 and 7. In Fig. 7, it can be noted that the level of the dissipation factor of the insulation at 90-per-cent relative humidity was still less than 0.10 even at the longest aging time, whereas with the organic varnish on the same base material, asbestos paper, the dissipation factor rose high above 0.5 even with relatively short aging times. The dielectric tests of Figs. 5 and 7 were made after 20 hours of humidification at 90-per-cent relative humidity.

It has been observed in a good many cases that the dielectric strength is reduced appreciably by high humidity only when the dissipation factor rises to a high value at the same time. The decreased dielectric strength is believed to be a result of dielectric heating. This effect will depend on the time of voltage application and the thickness of the insulation. Tests on humidified and aged resin materials with mineral fillers of a variety of sorts have all indicated that the impulse strength of the humidified material is not appreciably lower than the impulse strength of the same material dry. The 60-cycle strength, however, is much lower with humidification if the dissipation factor is high. This fact is important to remember when planning tests on humidified aged insulation. If overvoltages in practice are restricted to surges of short duration, then it would not be sensible to make overvoltage proof tests with 60-cycle voltages.

Summary

An attempt has been made to point out a number of factors affecting the results and the applicability of dielectric tests as a measure of insulation aging and life. The discussion has been directed particularly at a dielectric test on sheet insulation, but much of what has been said can indirectly apply to other aging tests on insulation. In its present form, the test can be regarded only as a measure of the relative life of insulating varnishes and not of the absolute life. It may possibly become a measure of the absolute life with particular base materials when satisfactory functional tests of these varnishes in apparatus or models are made to permit correlation. In the application it would seem necessary that something more than air dielectric strength be required as the function of the insulation to obtain a correlation with this test.

This discussion may make the problem of evaluating insulation life seem complicated but it is felt that, at this stage of the approach to the problem, it is better to recognize such variables than to ignore them and proceed rapidly to some wrong conclusions. In the end it seems inevitable that, if any positive result about the life of insulation materials is to be reached, some arbitrary decisions must be made, but these decisions should be made in the light of as much information as possible.

Vol.
Pt. 1
1955

References

1. Introduction to AIEE Standards. *AIEE Standard No. 1*, Feb. 1954.

2. Electrical Deterioration Treated as a Chemical Rate Phenomenon, T. W. Dakin. *AIEE Transactions*, vol. 67, pt. 1, 1948, pp. 113–18.

3. Aging of Dry-Tape Transformer Insulation, H. C. Stewart, L. C. Whitman. *AIEE Transactions*, vol. 67, pt. II, 1948, pp. 1600–07.

4. Aging of Class-B Insulating Material in Nitrogen, H. C. Stewart, L. C. Whitman. *AIEE Transactions*. vol. 70, pt. I, 1951, pp. 436–39.

5. Aging Evaluation of Dry-Tape Transformer Insulation Systems, H. C. Stewart, L. C. Whitman, A. L. Scheideler. *AIEE Transactions*, vol. 72, pt. III, Apr. 1953, pp. 267–77.

6. Calculation of Life Characteristics of Insulation, L. C. Whitman, Paul Doigan. *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 193–98.

7. The Deflected Beam—Film Rupture Test Applied to Sheet Insulation, H. J. Morgan, K. N. Mathes. *Special Technical Publication No. 161*, American Society for Testing Materials, Philadelphia, Pa., 1954, pp. 86–109.

8. A Method for Evaluation of the Thermal Aging Stability of Flexible Sheet Insulation, R. M. Piettner, C. G. Currin. *Ibid.*, pp. 111–20.

◆

Discussion

M. L. Manning (Pennsylvania Transformer Company, Canonsburg, Pa.): This paper merits consideration in that it points out the necessity for pausing to consider results and applicability of dielectric tests as a measure of insulation aging and life. The authors are to be congratulated on the open-door policy in making decisions in the light of as much information as possible. Summary papers like this are of much use in evaluating various test methods.

In the section "Factors Affecting the Life Test Results," the statement is made that electrode area is especially important when testing nonuniform material. In Fig. 3, a comparison of 1/4-inch and 2-inch diameter electrodes is made showing about 60 per cent reduction in dielectric strength for 2-inch diameter electrodes after 60 days' aging.

Electrode area is most important. To make the issue clear, why not use roller electrodes for flexible sheet insulation whereby the whole area of the sheet receives voltage stress? By roller electrode test, I mean the use of two rollers, $2^1/_2$ inches in diameter and about 2 inches less in width than the sheet, in which the bottom roller is grounded and voltage is applied to the upper roller. A sheet speed of 7 feet per minute has been found satisfactory. By this method, guesswork is removed in that no doubt exists about omitting weak portions of the sheet in which improper bonding of the resin to the glass fibers has occurred. Such a test before and after heat aging has been found most satisfactory for transformer application. Usually, about 20 per cent lower dielectric strength has been found for fibrous glass base resin bonded materials by roller tests, than by 2-inch diameter electrode tests. Weak portions of the sheet area often are not found by 2-inch diameter electrodes, even though the material is said to be homogeneous.

A transformer pancake winding offers severe voltage stresses on major high-voltage to low-voltage barrier insulation structures. Experience has demonstrated that high-temperature insulating materials in barriers have not failed even on overvoltage tests if the roller electrode test was used on component insulation.

T. W. Dakin, H. M. Philofsky, and **W. C. Divens:** As Mr. Manning indicates, we have indeed tried to look critically at the thermal aging test procedures which are developing rapidly and may in the future serve as a basis for the temperature rating of insulating materials. We should like to be sure that any new basis for rating or classifying materials is placed on as solid a foundation as reasonably possible. Mr. Manning has emphasized one of the factors that is the area of the test or the relative sample size. This is important in any test of insulating materials which are to be used in large areas in an apparatus. The complete area proof test referred to is certainly desirable in ensuring a satisfactory performance. But the damage of the insulation in a proof test where corona occurs should be guarded against. This has been shown to be a factor in some cases in using alternating voltages for proof testing.

In applying the results of any sample test to larger area applications the minimum expected value should be used. If the dispersion of results in the smaller sample test is obtained carefully, it is in principle possible to predict this expected minimum value with a calculable degree of confidence, without actually measuring it.

---

l, 74
1
55

Functional Evaluation of Magnet Wire Insulation

E. L. BRANCATO
MEMBER AIEE

R. S. PHILLIPS
ASSOCIATE MEMBER AIEE

Synopsis: A procedure developed at the Naval Research Laboratory to evaluate the thermal aging properties of magnet wire is presented. The method adopts a coil specimen and subjects it to such stresses as are experienced by the insulation system of a motor. The various factors which influence the aging properties of magnet wire are investigated. Aging data collected by this procedure are presented for a number of standard magnet wires.

THE past few years marked the development of a logical approach toward a functional evaluation of an insulating system.[1] The reference is representative of this tendency. This approach recognizes the necessity of imposing operational conditions upon any test method designed to measure expected insulation life. The Laboratory has long realized this necessity and, in fact, it has originated a method providing a functional evaluation of magnet wire.[2]

It is recognized that the magnet wire structures employed in electric equipments are generally exposed to a number of deleterious conditions. These conditions embrace temperature, voltage, humidity, contaminants, and mechanical stress. The latter stress results from vibration and from the expansion differentials of the system's components. It follows, then, that a functional evaluation of magnet wire must reflect the effects of all the operational conditions which serve to destroy, in time, the integrity of the insulation sheathing around the metallic conductor.

Coil Specimens

The specimen chosen consisted of a coil structure, Fig. 1, on a transite coil form *a* made by winding 100 feet of no. 26 American Wire Gauge insulated wire at a winding tension of 8 ounces. A 2-foot length of the same wire *b*, wound bifilarly with the coil wire at its physical center, is used to apply a dielectric stress to the insulation to detect insulation failures and, also, as a means of determining the temperature of the coil in the area of failure by resistance rise. Four-inch leads are brought from each of the two coils. At less than 1/8-inch distance from the coil form on each dielectric wire another 4-inch lead is attached. These wires serve as a current and potential leads for use with a Kelvin bridge. The main coil leads are used for heating.

Fig. 2 shows 20 identically wound coils rigidly mounted 1/4 inch apart on a 3/4-inch-diameter brass mandrel. This mass is wrapped with one layer of asbestos tape *a* which not only dampens any excessive vibrations set up in the coil leads *b* but also helps to minimize the temperature gradient caused by the difference between the coil temperature and the ambient temperature. The temperature gradient is reduced at each end of the mandrel by the addition of dummy asbestos tape coils *c*. The mandrel is mounted on a holder *d* which is constructed to make possible the removal of a group of coils from the test control circuit merely by disengaging several multijack connectors *e*. The holder makes it possible to solder all connections on the bench and to prepare another set of coils while the first set is on test. At the termination of a test, one holder can be immediately replaced by a second holder, permitting continuation of tests with minimum delay.

Test Procedure

Heating of Coil Specimens

The coils are self-heated by a 60-cycle-per-second current flowing through the 100-foot section. Unavoidably, temperature gradients exist over the coil section; however, the variation in temperature over the 2-foot length of dielectric wire is negligible. Since the observed failures are always between this wire and the heating coil, the aging temperature is measured by the resistance rise of the dielectric wire.

A bank of constant-potential transformers to minimize the normal line-voltage variations supplies the electric energy to heat the coil specimens. Temperature transients within the coils are further reduced by the temperature inertia of the brass mandrel upon which the coils are mounted. The coils are subjected to a constant ambient temperature

---

Paper 55-48, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing November 19, 1954.

E. L. Brancato and R. S. Phillips are with the Naval Research Laboratory, Washington, D. C.

The authors gratefully acknowledge the work contributed by A. T. McClinton and R. B. Whipple.

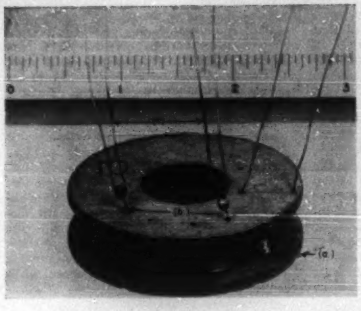


Fig. 1. Specimen coil

Fig. 2. Coil-mounting rack

of 60 degrees centigrade (C) provided by suitable ovens.

VOLTAGE STRESS

A voltage stress is established between the short dielectric wire and the heating coil with the application of 35 volts d-c, a maximum encountered in low-voltage circuit motors.[3]

HUMIDITY AND TEMPERATURE CYCLING

The heat-aging period is chosen to give the specimen an average life of 10 to 15 cycles. An aging cycle includes a series of exposures consisting of periods of heat aging, cooling to ambient, and humidification for approximately 48 hours in an ambient of 100-per-cent relative humidity at 35 C. A humidification period of 48 hours was found to be the minimum amount of time required for a complete moisture penetration of the coil.

VIBRATION OF COIL SPECIMEN

Even when aged insulation becomes brittle, it requires some form of mechanical stress to rupture the insulation film. Since electric equipments generally experience vibration during their operation, the coils are subjected to this form of mechanical stress. The magnitude of the vibration was fixed at 1.5 g's, which figure represents the average value obtained from a survey of a large number of 3,600-rpm motors.[2] The coil holders are therefore mounted on a vibration table during aging to provide this acceleration with a total displacement of 0.008 inch at 60 cycles per second.

NUMBER OF SPECIMENS

Those who study the properties of insulation systems know well that no single specimen can be considered representative of its class. Several coils apparently identically prepared and stressed will show an inherent spread in life about some mean point. The dispersion of failures about the mean point is caused by inhomogeneity of the dielectric materials, variable film thicknesses, and unavoidable differences in the preparation of the specimens. In addition to these factors, there are other statistically random effects such as those caused by impracticability of controlling temperature closer than ±2 C. Although it follows that the larger number of specimens the more representative the results will be, practical limitations restrict the number to 20 coils, a number which is felt to provide a sufficiently accurate statistical aging curve.

TEST CONTROL SYSTEM

A test panel, Fig. 3, provides individual coils with current control, dielectric stress, temperature measurement means (by the resistance rise method), and failure detection system. The dielectric stress is applied between the dielectric wire and the heating coil by a direct voltage source. Insulation failure is sensed by the rupture of a fuse in series with the dielectric wire. A recorder, which is in series with a circuit-sampling synchronous switch, indicates time of coil failure. The coils are sampled for failures periodically every half-hour or hour, depending on the expected life of the insulation.

Evaluation of Test Procedure

The majority of film-type insulations experience a definite sequence of effects when thermally aged. As the insulation begins to age, the volatile fractions of the dielectric material are driven off and the remaining material begins to embrittle. Concurrently with the process of embrittlement, shrinkage stresses usually are developed in the material. Eventually, the combined stresses caused by shrinkage and thermal expansion cause the embrittled film to crack. With additional aging the cracks become extended, thereby weakening the adhesion of the insulation to the wire. At this stage the vibration assumes a major role in the process of degradation. The action is one of loosening the cracked insulation

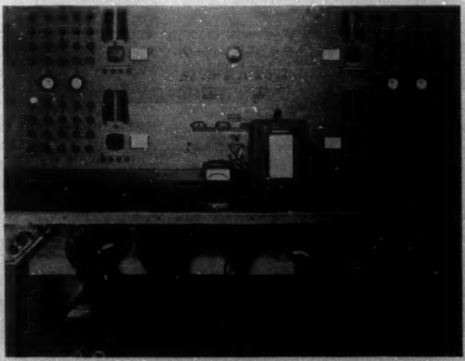


Fig. 3 Coil test control panel

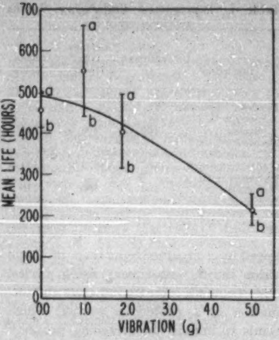


Fig. 4. Effect of vibration on life of Formvar-insulated coils at 160 C, 35-volt d-c stress. Lines a-b represent 95-per-cent confidence interval of the mean

and accelerating the flaking process. This procedure also helps to dislodge the loose insulation between conductors and, eventually, permits the establishment of metal-to-metal contact between them.

Because of the importance of vibration, a study was made to determine effects of vibration intensity on the life of insulation. Coil specimens were stressed at predetermined amounts of continuous vibration. The values chosen for test were zero, 1.0, 1.9, and 5 g's maximum acceleration at 60 cycles per second. A group of 20 coils of heavy Formvar insulation mounted on a test rack were aged at each value of vibration. The duty cycle consisted of 100 hours aging at 160 C followed by 64 hours of 100-per-cent relative humidity at 35 C. The dielectric stress for all specimens was 35 volts d-c.

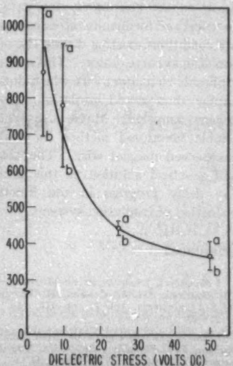


Fig. 5. Effect of voltage on life of Formvar-insulated coils at 160 C, 1.0-g vibration. Lines a-b represent 95-per-cent confidence interval of the mean

The results of this test, as seen in Fig. 4, reveal a decreased life with an increase in vibration. The effect is accentuated at the higher values of vibration. However, it should be recognized that vibration was not the only source of mechanical stress. The shrinkage-with-age and thermal expansion forces were sufficient to limit the life without the presence of induced vibration. If it were not for these latter forces, the life at zero vibration would be much greater.

Voltage stress is another factor influencing the life of insulation. A study of voltage effects was conducted on 4-coil assemblages, each consisting of 20 coils of heavy Formvar. The four groups of specimens were aged simultaneously at 160 C, but each group was stressed at different levels of voltage as follows: 4.5, 10, 25, and 50 volts d-c respectively. Each coil assemblage was aged at a duty cycle similar to the one described in the vibration tests and was subjected to a constant vibration of 1.0 g. The results shown in Fig. 5 reveal a strong dependence of insulation life on voltage stress. For example, increasing the dielectric stress from 10 to 25 volts causes a reduction of approximately 35 per cent in the mean value of insulation life.

EFFECT OF CYCLING AND HUMIDITY ON INSULATION LIFE

A spectrum of motor duty cycles ranges from very-short-time duty to continuous operation. The humidity exposures under these various conditions range from low to high. To determine the relative effects of the various conditions on the life of insulation, three types of operating procedures were chosen for investigation.

The first condition represented continuous-duty operation whereby the equipment is maintained at a rated temperature until failure occurs. No changes in temperature were permitted during these tests. The second condition, named cycling-without-humidity operation, consisted of a heat-aging period followed by a cooling period in which the specimens were allowed to reach room temperature at a relative humidity of less than 50 per cent.

The third condition, called cycling-with-humidity operation, consisted of a heat-aging period followed by a period of high humidity. After the heat-aging period, the samples were cooled to room temperature and placed in a chamber of 100-per-cent relative humidity for approximately 2 days. In choosing a duration for the heat-aging period, enough time was allowed to give the specimens an average life of 10 to 15 cycles of heat-aging and cooling.

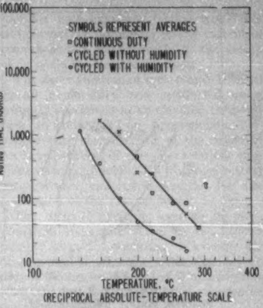


Fig. 6. Average life of Formvar coils

Aging tests were conducted on heavy Formvar insulation at the three conditions described in the foregoing. All specimens were subjected to identical conditions of vibration and voltage stress. The results of these tests are presented graphically in Fig. 6. It will be noted that very little difference in life was obtained between the continuous-duty and the cycling-without-humidity operations. However, the cycling-with-humidity tests revealed considerably less insulation life. This agrees with the generally accepted view that humidity is a sensitive probe for locating weaknesses in insulation. According to this view, corresponding insulation weakness must have occurred in the specimens subjected to continuous-duty and cycling-without-humidity operations; the absence of a searching agent delayed the sensing of failures until the insulation was sufficiently aged to flake off and allow the establishment of metal-to-metal contact. It is not excluded, however, that humidity is more than a probe and contributes to the deterioration.

COIL-AGING CYCLE

It becomes apparent from the results of these investigations that the factors contributing to the degradation of insulation must be incorporated in the procedure intended for an evaluation of the thermal-aging characteristics of magnet wire. These factors, in addition, must be introduced at such levels of magnitude most nearly simulating average operating conditions of electric equipment. With these thoughts in mind, a magnet-

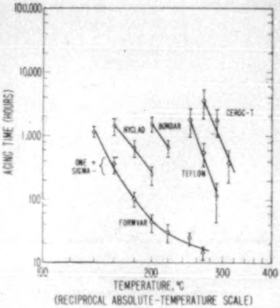


Fig. 7. Aging characteristics of magnet wire insulations

wire-aging technique was standardized to permit running 10 to 15 aging cycles to evaluate the average life. The aging cycle includes a thermal-aging phase wherein the specimens are concurrently subjected to 35-volt d-c electrical, and 1.5-*g* vibrational stresses. Following the thermal phase, the coil specimens are cooled to room temperature and, subsequently, exposed to a 100-per-cent relative humidity ambient for approximately 2 days. At the end of the humidity phase, the coil specimens are immediately returned to the thermal-aging phase and the aging cycle is repeated.

Aging Characteristics of Magnet Wire Insulations

Several types of magnet wire insulations were evaluated using the above aging cycle. These insulations were: Teflon, Ceroc-T, Bondar, and Nylclad. In addition, the effects of varnish on the aging properties of a Formvar type of magnet wire were determined.

When Teflon-insulated wire first became commercially available, an evaluation disclosed it to have very erratic characteristics. This behavior was attributed to poor bonding as was evidenced from the examination of specimens of unaged wire. Acquisition of better wire was made at a later date and consistent results were obtained. These results are presented in Fig. 7.

Ceroc-T insulation consists of an inner layer of ceramic material and two outer layers of Teflon. The ceramic material provides the insulation with high-temperature stability and prevents the pull through type of failure (typical for Teflon insulation). The Teflon cover protects the ceramic material from abrasion and reduces moisture absorption. A magnet wire with this combination of materials was evaluated in coil form and was found to possess life that is many times greater than Teflon alone.

For comparison, data are also presented in Fig. 7 for the newer types of magnet wire insulation, namely Bondar and Nylclad.

Magnet wire is seldom used without the addition of a varnish which adds greatly to the reliability of the electric equipment. The effect of varnish on the life of a Formvar insulation was determined on two sets of coils which were dipped in commercial varnishes and baked, then aged at 160 C according to standard procedure. One set was impregnated with a class-*A* clear baking varnish and the other set was treated with a silicone varnish. The aging results are shown in Table I and indicate that the organic varnish doubled the life of the insulation. The silicone varnish had no effect on the life of the Formvar, thereby making this more expensive varnish unsuitable for use with certain class-*A* materials.

Problem of Glass-Served Magnet Wire

Class *B*, glass-served magnet wire was also aged at various temperatures, giving erratic results. When operating at temperatures ranging from 220 to 290 C, the first coil failures usually occurred at approximately 1,000 hours regardless of the temperature. The subsequent failures were very erratic, with the majority of coils experiencing very long life. This erratic nature is inherent in that glass fibers serve as an effective separator, which breaks down only under certain unusual conditions to establish metal-to-metal contact. It is apparent that actual motors with glass-served magnet wire usually experience failures resulting from surface contaminants rather than from the embrittlement and flaking procedure occurring in film-type magnet wire. A number of commercial laboratories have experienced the same difficulty on motors and, consequently, they have established as a criterion of failure, a marked drop in the "wet" insulation resistance. The authors' experience with magnet wire specimens has not yet proved this criterion to be successful. It has been found that the marked drop in resistance occurs too soon after aging begins, resulting in a pessimistic value of insulation life. In a number of cases the drop in the wet insulation resistance was followed by a progressive increase in ohmic value during subsequent aging cycles. The resistance increase indicates the formation of a protective coating which tends to inhibit the searching property of the voltage stress.

Table I. Impregnated Coil-Aging Results

Insulation	Temperature, C	Average Life, Hours	Sigma, Hours
Formvar + General Electric *9574* varnish	160	800	105
Formvar + Dow Corning *997* silicone varnish	160	372	114
Unimpregnated Formvar	160	360	110

The coil-aging technique does not provide for contaminants of sufficient quantity to bridge the gaps of the glass-insulated conductors; therefore this aging technique must be considered unsuitable for evaluating these types of insulation. It is apparent that a different criterion of failure for glass insulation must be developed.

Summary

In this paper the effects of temperature, voltage, humidity, and mechanical stress upon the aging properties of magnet wire are described. It is observed that humidity plays a dominant role influencing the aging characteristic. However, this characteristic is also a sensitive function of voltage and vibration stress. A method of evaluating magnet wire has been adopted which places special emphasis on the continuous application of electrical and mechanical stresses to simulate conditions existing during the operation of an average motor. This method is applicable to magnet wire which does not employ glass-served insulation. To the authors' knowledge there is, as yet, no reliable functional method to evaluate glass-served magnet wire. The absence of a practical solution to this problem may delay progress in the functional evaluation of insulation systems.

References

1. FUNCTIONAL EVALUATION OF MOTOR INSULATION SYSTEMS, **G. A. Cypher, R. Harrington.** *AIEE Transactions*, vol. 71, pt. III, Jan. 1952, pp. 251–53.

2. AGING CHARACTERISTICS OF ELECTRICAL INSULATION, **R. E. Whipple.** *Report No. 3708*, Naval Research Laboratory, Washington, D. C., July 24, 1950.

3. DESIGN OF ELECTRICAL APPARATUS (book), **J. H. Kuhlmann.** John Wiley and Sons, Inc., New York, N. Y., 1946, p. 47.

Discussion

L. E. Sieffert and H. P. Walker (United States Navy Department, Washington, D. C.): The method presented appears to cover similar ground to that discussed in the paper by Currin and Dexter.[1] One may possibly supplement or replace the other. It is highly desirable that additional work be undertaken to determine the extent to which these methods may be employed, and the results of each method correlated with studies on motorettes and motors. To determine the place that each of these test methods may have in final evaluation as a component of a system, it would appear that much additional test data, obtained under the varying conditions at different laboratories, must be studied. These should include varnished treatment of coils.

Reference

1. A Method for Evaluation of Thermal Stability of Magnet Wire Enamel, C. G. Currin, J. F. Dexter. *AIEE Transactions*, vol. 74, pt. I, May 1955, pp. 227–32.

William T. Gordon (Westinghouse Electric Corporation, Lima, Ohio): The test procedure described seems quite adequate for the singular purpose of the evaluation of magnet wire insulation under conditions similar to those experienced in a motor. I believe that the evaluation of the separate components of an insulation system is often a necessary supplement to the evaluation of the complete system. In the final analysis, however, it is the ability of all components to function as a system which determines their suitability as motor insulation. Accordingly, the importance of testing complete insulation systems as described in the AIEE Standard must be recognized.[1]

The paper states that the heat-aging period was chosen to give the specimen 10 to 15 cycles. How is this accomplished? In the course of testing a number of samples have any survived 15 cycles or 1,500 hours of heat aging? Are the dummy coils removed during the humidification phase of a cycle?

Table I is very interesting. The results recorded clearly show the importance of proper selection of impregnating varnishes for use in an insulation system. One explanation for the failure of the silicone varnish to improve the insulation life of the Formvar-wound coils might be that the deleterious effects of the varnish-curing temperature on the enamel film offsets the gain in life to be expected from the varnish.

It is the independent or collective efforts of the manufacturers and users of electrical equipment such as those described in the paper which will go far toward making a science of the application of insulation.

Reference

1. Test Code for Evaluation of Systems of Insulating Materials for Random-Wound Electric-Machinery. *AIEE Standard No. 1C*, Jan. 1954.

E. L. Brancato and R. S. Phillips: Mr. Sieffert's and Mr. Walker's comments bring into proper perspective the spirit in which the paper was presented. The coil-testing procedure involves a number of complexities not inherent in the method described by Mr. Dexter in his report (see reference 1 of Mr. Sieffert's and Mr. Walker's discussion). It is our conviction that the technique employing twist specimens is an excellent and simple method for obtaining a measure of the relative thermal stability of magnet wire while the coil testing procedure provides a prediction of insulation life as may be obtained in a magnet wire structure of a system. More work is being done at the Naval Research Laboratory to compare the two methods but it is hoped that other members will obtain experience and data to permit a more thorough comparative evaluation of the methods.

Appreciation is expressed for the encouraging remarks of Mr. Gordon. We cannot help but concur with his conclusion that the real measure of the quality of an insulating component must be judged ultimately by its behavior in a complete system. It is the hope of the authors that the acquisition of a more comprehensive knowledge of the mechanisms of deterioration in insulation together with improved testing techniques will permit the functional test of individual components to help directly in the design of an optimum insulation system.

The determination of cycle length is based on an exploratory test conducted on a group of specimens at an elevated temperature to establish an approximate life-temperature characteristic of the material. An extrapolation of this characteristic provides sufficient information to choose testing temperatures and cycle lengths for the evaluation of the magnet wire.

The 10 to 15 cycles to average life mentioned in the paper refers to the desired number of cycles occurring to average life of a group of 20 coils. The first failure usually occurs at the end of 5 aging cycles while the last failure may have endured as many as 20 cycles.

The construction of the coil-mounting rack does not allow the removal of the dummy coils or the asbestos tape covering during the humidification phase of the cycle. We had discovered, by insulation resistance measurements, that the porous tape does not act as a very effective moisture barrier and permits absorption of moisture by the coils. However, the tape will prevent the formation of visible condensate on the insulation surface.

The possibility that the baking of the silicone varnish will cause sufficient initial aging of the Formvar magnet wire to reduce the operating life below that obtained with conventional varnish is an interesting one. During the analysis of the resulting data, this possibility was considered and it was concluded that the baking of the varnish contributed negligible aging. The curing of the silicone varnish was done at 150 C for a total period of 35 hours. If the 10-degree aging rule is employed, this time may be converted to an equivalent time of 17.5 hours at 160 C. The addition of this time to the hours life, indicated in Table I, will give an approximate total of 390 hours of life for Formvar magnet wire impregnated with silicone varnish. This value is still significantly below the corresponding value of 800 hours obtained for Formvar magnet wire aged at the same temperature, but impregnated with a compatible varnish.

An All-Magnetic Audio Amplifier System

J. J. SUOZZI
ASSOCIATE MEMBER AIEE

E. T. HOOPER
ASSOCIATE MEMBER AIEE

ALTHOUGH magnetic amplifiers have been used as audio amplifiers, the necessity of generating the high carrier frequency has made their use impractical. The objection to the use of rotating machinery to produce this carrier is the mechanical problem introduced by moving parts, while the use of the vacuum-tube oscillator as a high-frequency source introduces the disadvantages of fragility, short life, and increased maintenance problems. Although switching transistors associated with saturable reactors can be used to produce a high-frequency carrier, problems of performance and reliability arise because of the early stage of development.

Since the primary advantages of the magnetic amplifier are its ruggedness, reliability, and static nature, it would be advantageous to use it with a high-frequency power supply of essentially the same nature. Such a carrier source is the static frequency multiplier consisting of transformerlike devices and capacitors. However, in the past when these have been used, the resulting unit was prohibitively large and introduced much distortion into the system. The combination of static multipliers of improved

---

Paper 55-70, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted September 24, 1954; made available for printing November 19, 1954.

J. J. Suozzi and E. T. Hooper are with the U. S. Naval Ordnance Laboratory, Silver Spring, Md.

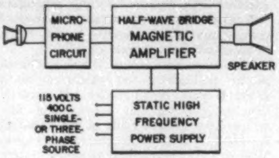


Fig. 1. Block diagram of the magnetic audio amplifier system

design with the half-wave bridge-type magnetic amplifier[1,2] has led to the development of an audio system of practical size, low distortion, and good bandwidth.

The High-Frequency Power Supply

The audio system consists of a static, high-frequency power supply, a half-wave bridge-type magnetic amplifier, and suitable input and output circuits working from a 400-cycle-per-second (cps) 115-volt single- or 3-phase source. A block diagram of the system is illustrated in Fig. 1.

The high-frequency power supply consists of cascaded frequency triplers of the magnetic core type. This type of frequency multiplier has been in use for some time. In 1913, a single-phase transformer-type frequency tripler was described by Joly.[3] Later, the single-phase bridge-type frequency tripler was developed by Von Bronk[4] in Germany. A polyphase magnetic frequency multiplier was introduced by Huge[5] in 1947 and other variations of polyphase multipliers were shown by Alexanderson et al[6] in 1948 and Logan[7] in 1950. Further discussions of the use and cascading of frequency multipliers have been made by Harriott,[8] Geyger,[9] and Johnson and Rauch.[10]

From this wealth of information, work was begun on the development of a static high-frequency source to drive a magnetic audio amplifier. A primary consideration was to minimize the physical size while producing sufficient power with a good sinusoidal wave form. Consequently, a carrier of 10,800 cps was chosen since this would entail the least amount of frequency multiplication while allowing a sufficient bandwidth to be obtained when using a half-wave magnetic amplifier. Obviously higher fidelity can be obtained by a further increase in the carrier frequency.

To obtain this carrier from the 400-cps mains, three frequency triplers were cascaded. All of the three basic types of triplers with modifications and design changes were used in various combinations. Figs. 2 and 3 show the circuit diagrams of each.

Up to the present time there have been two general methods of operating the 3-phase tripler. One method[5] utilizes only saturable reactors in each of the three phases while taking the output across a load in the neutral lead; see Fig 2(A). This output wave form appears, as shown in Fig 2(E), when a sharply saturating material with a rectangular hysteresis loop is used. The other method,[6] Fig. 2(B), utilizes a linear reactor and capacitor combination with the saturable reactor in each phase to produce third and higher harmonics. It has been shown that the use of a linear reactor alone in series with the saturable reactor is desirable for wave-shaping purposes; see Fig. 2(C). However, the addition of these three linear reactors increases the size of the unit considerably. To obtain the advantages of this method of wave-shaping without excessive increase in size, the authors have used only one linear reactor with three separate windings, as shown in Fig. 2(D). This improvement is possible because of the switching action of the saturable reactors. Since in this 3-phase system only one saturable reactor will be saturated at any instant of time, the linear reactor is effectively in only one phase at a time. The output of the tripler can be fed into a tuned circuit to obtain a better sinusoidal wave form. In this case, only a capacitor is used for tuning.

The 3-phase tripler is highly efficient. Efficiencies (watts out divided by watts in) as high as 98 per cent have been obtained. The 3-phase tripler used in the power supply for the audio system has an efficiency of 90 per cent.

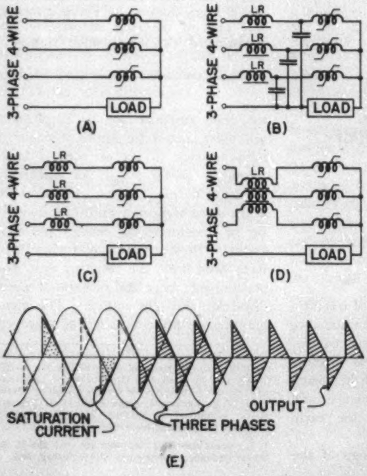
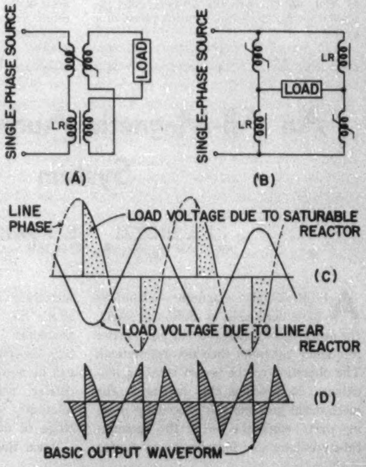


Fig. 2 (left). Circuits and wave forms of the 3-phase tripler

A—Basic circuit with saturable reactors only
B—With added inductance - capacitance circuits for production of higher harmonics
C—With linear reactors for wave-shaping
D—With single linear reactor for wave-shaping
E—Basic output wave form before tuning or wave-shaping

Fig. 3 (below). Circuits and wave forms of the single-phase tripler

A—Transformer-type single-phase tripler
B—Bridge-type single-phase tripler
C—Superposition of the sinusoidal and distorted wave forms
D—Basic output wave form before tuning



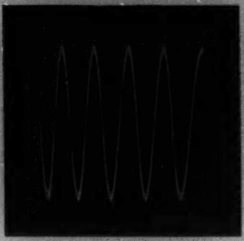


Fig. 4. Output wave form of 10,800-cps power supply

The operation of the single-phase triplers depends on the superposition of a sinusoidal and distorted voltage or current. A linear reactor is used to produce the sinusoidal voltage and necessary phase shift while a saturable reactor produces the desired distorted wave form. The superposition of the two is accomplished by the transformer-type circuit,[3] Fig. 3(A), or the bridge-type circuit,[4] Fig. 3(B). Fig. 3(C) shows the two wave forms contributed by the action of the linear and saturable reactors while Fig. 3(D) shows the basic output wave form which is a summation of the two. This output is rich in third harmonic and by means of appropriate tuned circuits can be made sinusoidal.

l, 74
1
55

Simple and improved design techniques have been developed to obtain maximum efficiency with a minimum of size. Efficiencies (watts out divided by watts in) of 85 per cent have been obtained with the single-phase tripler. The efficiency of each single-phase tripler used in the power supply of the audio system was 70 per cent. However, to improve the wave form, a sacrifice in power was made in the third tripler reducing its efficiency to 35 per cent.

Consequently, the high-frequency power supply, consisting of a 3-phase tripler followed by two single-phase triplers, used in the audio system has an over-all efficiency of 21 per cent which compares favorably with many vacuum-tube power supplies. Efficiencies almost as high can be obtained when three single-phase triplers are cascaded.

The output of the high-frequency power supply is 25 watts at 40 volts of 10,800 cps with a harmonic distortion of less than 5 per cent. An oscillogram of this output wave form is shown in Fig. 4. Double this power output can be obtained with some increase in harmonic distortion, as explained in the foregoing. This output has been achieved in a package size of only 4 by 5 by 6 inches.

The Magnetic Audio Amplifier

The magnetic amplifier designed for operation with the high-frequency power supply must possess certain characteristics. Its power demand must be low since limited power is available, and overdriving the carrier source will produce distortion and spurious frequencies in the audio output. Also, the amplifier should draw little or no direct current from the source since this too disturbs the operation of the frequency multiplier and produced distortion. A further requirement for the amplifier is a bandwidth sufficiently high to reproduce satisfactorily audio signals.

To meet the foregoing requirements, a 2-stage half-wave magnetic amplifier[1,2] was chosen. Each stage consists of two saturable reactors, appropriate selenium rectifiers, and other static components. The power drain was kept below 25 watts while the two stages were designed symmetrically to draw approximately the same amount of direct current, thus appearing as an a-c load to the power supply. The a-c power gain of the amplifier is approximately 1,000 while the physical size including the microphone circuit components is $1\frac{1}{2}$ by 5 by 4 inches.

The transfer function of the half-wave magnetic amplifier includes a dead-time function ϵ^{-st}, where t is the time delay. The dead time is $1\frac{1}{2}$ cycles of the carrier frequency for the 2-stage amplifier. This produces a linearly increasing phase distortion which is of no consequence in the audio system since the ear cannot detect the phase angles of the various frequency components.

The complicated nature of control in the magnetic amplifier is such as to cause both frequency and nonlinear distortion under certain conditions. These effects are apparent upon referring to the wave forms of Fig. 5. By nature the magnetic amplifier is controlled during some presetting period. In this case, this period occurs during alternate half-cycles of the carrier. The portion of the signal that is effective in controlling the amplifier is phase-reversible, the phase and magnitude of the output on each half-cycle is determined by the phase and magnitude of the effective control half-cycle. For signal frequencies below half the carrier frequency, e.g., 3,600 cps as shown in Fig. 5(A), the output is essentially 3,600 cps. At half the carrier frequency, i.e., 5,400 cps, the output is either 5,400 cps, as in Fig. 5(B), or zero, as in Fig. 5(C), depending on the phase of the signal with reference to the carrier. For signal frequencies greater than half and less than the carrier frequency, e.g., 7,200 cps, the output is seen to contain a large component of 3,600 cps; see Fig. 5(C). This beat frequency between the signal and carrier frequencies is the source of the nonlinear distortion mentioned in the foregoing. In addition, for signal frequencies approaching half the carrier frequency, reduced output is obtained depending on the relative phase between the signal and the carrier.

Despite these adverse effects caused by the nature of the amplifier, the over-all frequency and nonlinear distortion of the amplifier working from the frequency multiplier is less than 10 per cent over the bandwidth of the system. This is exclusive of the carrier which was suppressed by an appropriate filter. A plot of distortion measured on the amplifier working from its magnetic power supply is shown in Fig. 6(A).

The bandwidth (3 decibels down) of the

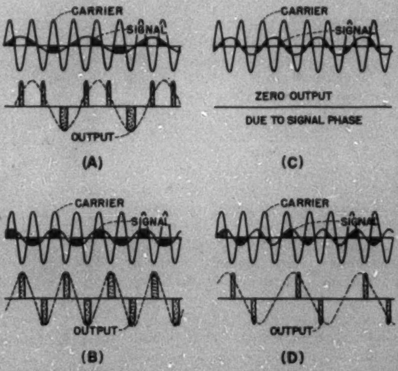


Fig. 5. Control mechanism of the half-wave phase-reversible magnetic amplifier

A—With signal frequency of one-third the carrier frequency
B—With signal frequency of one-half the carrier frequency and maximum output
C—With signal frequency of one-half the carrier frequency and zero output
D—With signal frequency of two-thirds the carrier frequency

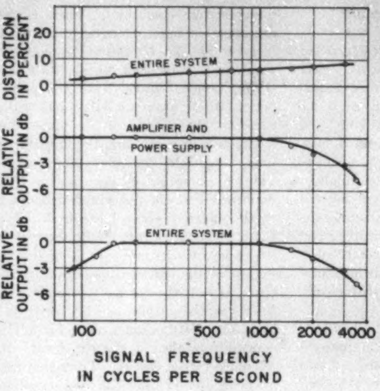


Fig. 6 (left). Performance curves of the magnetic audio amplifier system

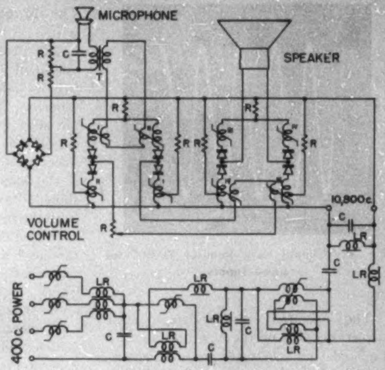


Fig. 7 (right). Complete schematic diagram of the magnetic audio amplifier system

magnetic audio amplifier is zero to 3,000 cps, as indicated in Fig. 6(B). The upper limit of this bandwidth is determined by the detrimental phase shift discussed, and stray capacitance in the amplifier. It is to be noted that the lower limit of the bandwidth goes to zero in that the half-wave magnetic amplifier functions as a d-c amplifier.

Performance of the Audio System

The schematic diagram of the entire audio system is shown in Fig. 7. The amplifier is driven by a carbon microphone and suitable transformer. The direct current for the microphone circuit is obtained by rectifying and filtering a portion of the 10,800-cps voltage. The output of the amplifier is fed directly into the voice coil of a permanent magnet speaker. The physical size of the entire unit can be seen in Fig. 8. The amplifier and power supply can be combined in a package 5 by 5 by 6 inches.

The bandwidth of the entire system is 90 to 3,000 cps, as seen in Fig. 6(C). The fall off below 90 cps is attributable to the input circuit and the frequency response of the speaker.

The audio output power is 2.5 watts with less than 10-per-cent distortion over the frequency range. This system possesses all the advantages of magnetic amplifiers, i.e., reliability, ruggedness, long life, instant warm up, and a usable temperature range limited only by the selenium rectifiers. Because of the instant warm-up characteristic, the system will operate with no stand-by power since the unit need be excited only when voice amplification is desired. Thus, a push-to-talk switch could be used as an on-off power switch. This would result in the complete elimination of quiescent power and the attendant heat dissipation problem.

A further by-product of this characteristic is the possibility of derating the frequency multiplier and amplifier with a resultant decrease in size when the duty cycle permits. The unit described in this paper was designed for 100-per-cent duty cycle, a condition not often encountered. With the additional advantages of potting the entire amplifier and power supply, an audio system of complete dependability and long life can now be realized.

Conclusions

The construction of an all-magnetic audio amplifier system is entirely feasible both in size and performance. In ruggedness, dependability, and long life, such a unit should excel all others presently known in the art. At present, this type of audio system is suitable for intercommunication, facsimile reproduction, public address, and battle command systems. As higher source frequencies become available, this solution to the audio amplifier problem will find wider application.

Fig. 8. Complete magnetic audio amplifier system

References

1. An Improved Magnetic Servo Amplifier, C. W. Lufcy, A. E. Schmid, P. W. Barnhart. *AIEE Transactions*, vol. 71, pt. I, 1952, pp. 281–89.

2. Design Considerations of the Half-Wave Bridge Magnetic Amplifier, C. W. Lufcy, H. H. Woodson. *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 220–26.

3. Static Frequency-Transformer, M. Joly. *U. S. Patent 1,118,935*, Dec. 1, 1914.

4. Frequency Multiplier, O. Von Bronk. *U. S. Patent 1,678,965*, July 31, 1928.

5. Frequency Multiplier, H. M. Huge. *U. S. Patent 2,424,237*, July 22, 1947.

6. Electric Frequency Transformation System, E. F. W. Alexanderson, A. H. Mittag, M. W. Sims. *U. S. Patent 2,451,189*, Oct. 12, 1948.

7. Means for Increasing Frequency, F. G. Logan. *U. S. Patent 2,517,396*, Aug. 1, 1950.

8. Magnetic Frequency Converters, L. C. Harriott. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 9, 1953, pp. 78–87.

9. Magnetic-Amplifier Circuits (book), W. A. Geyger. McGraw-Hill Book Company, Inc., New York, N. Y., 1954, pp. 249–51.

10. Magnetic Frequency Multipliers, L. J. Johnson, S. E. Rauch. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 448–52.

◆

Discussion

E. Lehtonen (Sperry Gyroscope Company, Great Neck, N. Y.): The single 3-winding reactor shown in Fig. 2(D) does indeed offer some economy of material compared to the three separate reactors shown in Fig. 2(C). However, the 3-winding reactor is effectively a single-winding reactor since the windings are presumably tightly coupled. This point is more clearly seen in Fig. 7 where the three reactor windings are shown with a common connection. The voltages at the left-hand end of the windings are effectively at the same potential, neglecting current-resistance drops. If the left-hand ends were joined together, either lower losses could be achieved or a smaller reactor could be used.

This parallel-connected reactor corresponds to the reactor L_s in Figs. 1, 2, and 3 of the paper by Smith and Salihi.[1]

Reference

1. Analysis and Design of a Magnetic Frequency Multiplier, O. J. M. Smith, J. T. Salihi. *AIEE Transactions*, vol. 74, pt. I, March 1955, pp. 99–106.

J. J. Suozzi and E. T. Hooper: We wish to thank Mr. Lehtonen for his very worthwhile observation. We should, however, like to add the following comment.

When a separate reactor is used in each phase, as in Fig. 2(C), both wave shaping and protection from excessive circulating currents are obtained should the 3-phase source become umbalanced and cause two saturable reactors to be saturated simultaneously.

When using only a single-winding reactor in the neutral lead, such as L_s in Figs. 1, 2, and 3 of the paper by Smith and Salihi, only wave shaping is obtained.

Considering the 3-winding reactor in Fig. 2(D), it is seen that the separate reactors are a special case of this configuration, i.e., zero coupling. Of course, zero coupling can be obtained by using separate reactors. It is further seen that the single-winding reactor is also a special case of the 3-winding configuration, i.e., essentially perfect coupling, as Mr. Lehtonen points out.

In addition, with coupling between zero and perfect, any desired combination of characteristics can be obtained. Therefore, utilizing this configuration with less than perfect coupling, both wave shaping and protection against circulating currents can be obtained. The size, of course, depends on the amount of coupling and the demands of the circuit.

Empirical Equations for the 60-Cycle Spark-over Gradients in Commercial Oil

W. F. GAUSTER
MEMBER AIEE

INTENSIVE experimental work on dielectric phenomena in liquids has been done during the last 40 years. There is considerable interest on the engineer's part in obtaining reliable information about the dielectric properties of liquids, especially of commercial oil, used as insulation and as a cooling medium for transformers, circuit breakers and much other electrical apparatus. Many valuable experimental results have been obtained for dielectric strength, corona, dielectric losses and other dielectric properties of commercial oil.[1–5] In connection with the first-mentioned problem, that of the dielectric strength or oil, numerous details have been carefully studied: for example, the influence of electrode geometry, the rate of voltage rise, humidity, dissolved gases and different kinds of contamination. However, as Whitehead has mentioned,[6] "Breakdown in liquids could readily become a compendium of unrelated observations" because very little was known about the physical background of this field.

Only recently, experimental technique has been so much improved that reproducible results for the impulse breakdown strength of carefully purified liquids for small electrode spacings could be obtained.[7,8] Physical models have been used to explain the dependence of the measured dielectric strength of liquids on electrode spacing[9] and on the molecular structure.[10,11] Furthermore, experimental work has been done on impulse corona in commercial oil,[12] which seems to be a good start for studying the physical phenomena involved in electrical breakdown of oil.

However, these results cannot yet be applied for finding numerical values of 60-cycle spark-over gradients for different electrode geometries in commercial oil. Therefore, it is interesting to attempt to interpret basic empirical data on the dielectric strength of commercial oil from an entirely different point of view. Several applications of "weakest link" probability theory to insulation variability have been made.[13,14]

W. R. Wilson developed a theory of the equivalent volume[15] and applied it to different electrode geometries. Unfortunately, as shown in this paper, this application of the theory of probability to spark-over in commercial oil does not lead to the desired results. It still seems necessary to use purely experimental data and to consider carefully appropriate empirical equations until a more general presentation of such relationships is achieved.

Empirical Equations for Oil Spark-over Gradients for Sphere Gaps of Various Radii

The results of an extensive series of experiments on the spark-over voltages of sphere gaps in commercial oil, performed with 60-cycle alternating voltage, are contained in Table LXIV (page 220) in reference 1. Sphere radii R from 0.159 to 6.25 centimeters (cm), and spacings x from 0.129 to 5.08 cm were used; and values for the spark-over voltage V and for the maximum spark-over gradient G were listed. The average spark-over gradient in the gap, measured along the axis connecting centers of the spheres, is

$$G_{av}=\frac{V}{x} \qquad (1)$$

The maximum spark-over gradient, which occurs at the intersections of the axis with the sphere surfaces, is

$$G=fG_{av}=f\frac{V}{x} \qquad (2)$$

The factor f is a function of the ratio ξ of the spacing x to the radius R.[16] In this paper only the case of symmetrical potential distribution (neither sphere grounded) is considered. With one sphere grounded, a different factor f_1 must be used.

Tables for the numerical values of $f(\xi)$ have been calculated.[17] The factor f does not consider any field distortion resulting from space charge. Peek states that for each value of the radius R the spark-over gradient G reaches a minimum value G_{min}, which corresponds to a horizontal (constant) part of the breakdown curve at sufficiently large spacing x. In Fig. 182 and Table LXVIII he listed these minimum spark-over gradients G_{min} and shows that these data can be expressed by an empirical equation

$$\{G\}_{min}=28.3\left(1+\frac{4}{\sqrt{R}}\right)\text{kv}_{max}/\text{cm} \qquad (3)$$

Paper 55-65, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 18, 1954; made available for printing November 30, 1954.

W. F. Gauster is with North Carolina State College, Raleigh, N. C.

The author wishes to acknowledge the valuable comments of J. Hagenguth, T. W. Liao, and K. H. Weber, and the assistance of D. S. Burton; and to thank Dr. A. Roth for providing supplementary data for the experiments performed by the Brown Boveri Company, Baden, Switzerland.

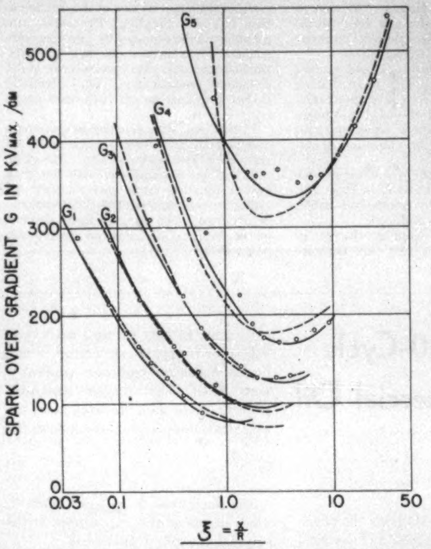


Fig. 1. Sixty-cycle spark-over gradients G_k of sphere gaps in oil for different radii R_k and spacings x. Encircled points are Peek's empirical data. Solid curves are best fitting parabolas, dashed curves are calculated with equations 9 and 10

with an error not greater than ±10 per cent (%). Calculated values are expressed by enclosing the letter between brackets { } or [ ]. Observed values are indicated by the letter without brackets.

To discuss Peek's experimental results for sphere gaps in oil in greater detail, a different physical parameter in the plot of the breakdown gradients will be used, which has often been applied to sphere gaps in gases.[18,19] This is shown in Fig. 1 where the spark-over gradients are plotted against $\log x/R = \log \xi$ (here log means logarithm to the base 10), instead of x as in Peek's book. The different points shown represent the experimental data from Peek's table LXIV. For each value $R_k = R_1, R_2 \ldots R_5$ of the radius R a smooth solid curve is drawn. They can be represented by equations of the form

$$\{G\} = \{G_{\min}\}(R) + N(R)\left[\log \frac{\xi}{\xi_0(R)}\right]^2 \qquad (4)$$

If ξ is plotted on a logarithmic scale, the curves are parabolas which are symmetrical about the vertical lines drawn through the vertexes $G_{\min}$ with the abscissas $\log \xi_0$. If a linear scale for ξ is used, the curves are very steep for small values of ξ, and become more and more flat for large values of ξ. Curves of this kind, observed for spark-over gradients in air, are shown by Whitehead.[20] It must be remembered that the ratio $G/G_{av} = f(\xi)$ used here is defined for symmetrical spheres infinitely far from ground, a condition, of course, never fully realized in practice. The numberical values of G for small values of R and large values of ξ should be used with caution because in these cases the connecting leads and the proximity of the tank walls will distort the electrostatic field of the sphere gap.[17]

The minimum values of the spark-over gradients $G_{1\,\min}$, $G_{2\,\min}$, and $G_{4\,\min}$ plotted in Fig. 1. are taken from Peek's experimental data as listed in his table L XVIII. $G_{3\,\min}$ and $G_{5\,\min}$ are between the experimental values and those calculated with Peek's equation 3. The values $N(R_k)$ and $\xi(R_k)$ are chosen so that equation 4 fits the experimental data for G as closely as possible. This was done by trial and error, because an evaluation with the use of the principal of least squares is too involved here. The following empirical equations have been used and the results are listed in Table I.

$$\left.\begin{aligned}
R_1 &= 6.25\ \text{cm};\ \{G_1\} = 82 + 90\left(\log \frac{\xi}{1.3}\right)^2\\
R_2 &= 3.12\ \text{cm};\ \{G_2\} = 98 + 100\left(\log \frac{\xi}{2.2}\right)^2\\
R_3 &= 1.27\ \text{cm};\ \{G_3\} = 125 + 110\left(\log \frac{\xi}{3.5}\right)^2\\
R_4 &= 0.555\ \text{cm};\ \{G_4\} = 169 + 170\left(\log \frac{\xi}{4.0}\right)^2\\
R_5 &= 0.159\ \text{cm};\ \{G_5\} = 335 + 210\left(\log \frac{\xi}{3.8}\right)^2
\end{aligned}\right\} \qquad (4A)$$

These equations represent the set of Peek's empirical data with deviations of less than ±10%. Obviously, this representation is not unique, since the coefficients $N(R_k)$ and $\xi_0(R_k)$ could be assumed in an appropriate way with values somewhat different from those shown in the foregoing and, nevertheless, a satisfactory approximation of the experimental data would result. The next goal should be to find simple empirical forms for the coefficients $N(R_k)$ and $\xi_0(R_k)$, similar to that in equation 3 for $\{G_{\min}\}$. For example, in the case of sphere gaps in air it is known that the spacing x_0, below which the apparent strength G begins to increase, is represented by the equation[21]

$$x_0 = 0.54\sqrt{R} \qquad (5)$$

Therefore, for air

$$\xi_0(R_k) = \frac{x_0}{R_k} = \frac{0.54}{\sqrt{R_k}} \qquad (6)$$

If it were possible to find similar empirical forms for $\xi_0(R)$ and $N(R)$ for sphere gaps in oil, then equation 4 alone would be sufficient for calculating the spark-over gradient for any value of R and ξ. However the empirical data available up to now for sphere gaps in commercial oil are inadequate to provide conclusive results when the foregoing method is used. Therefore an alternate method was developed as in the following.

Empirical Equation Representing Peek's Complete Set of Data for Sphere Gaps in Oil

It is possible with sufficient accuracy to find one single empirical equation which represents the complete set of Peek's data for sphere gaps in oil if we proceed in the following way: From equation 3 it follows that

$$(\{G_{\min}\} - 28.3)\sqrt{R} = 113.2 \qquad (7)$$

is a constant. A numerical evaluation of the term $(G - 28.3)\sqrt{R}$, where G now means the experimental value of the spark-over gradient for any value of R and ξ, shows that this term is a function of ξ alone (a similar empirical law was used by Whitehead for sphere gaps in air.[22]) to a fairly good approximation

$$(G - 29.3)\sqrt{R} = F(\xi) \qquad (8)$$

All the points contained in Peek's Table LXIV are considered in Fig. 2. A best-fitted curve is then drawn for $F(\xi)$. The vertex of this curve has the co-ordinates log 2.5 and 113 kv$_{\max}$/cm. The left and right branches of the curve can be approximated by parabolas of different slope, expressed as

$$\left.\begin{aligned}F(\xi)&=113+165\left(\log\frac{\xi}{2.5}\right)^2\ldots\xi\leq 2.5\\ F(\xi)&=113+\ 70\left(\log\frac{\xi}{2.5}\right)^2\ldots\xi\geq 2.5\end{aligned}\right\}\qquad(9)$$

If $F(\xi)$ is known, G is found by means of equation 8, put in the form

$$[G]=28.3+\frac{1}{\sqrt{R}}F(\xi)\qquad(10)$$

With the empirical equation 10 it is possible to compute the spark-over gradient with good approximation for any value of R and x contained in Peek's Table LXIV. These values are designed by $[G_k]$ and are listed in Table I. The results are not as accurate as those obtained by applying equations 4, but the accuracy seems to be sufficient if the relatively poor reliability of the experimental data is considered. In Fig. 1 the values $[G_k]$ computed with equations 9 and 10 are shown by means of dotted curves. The minimum values $[G_k]_{\min}$, obtained with these equations, follow exactly the values $\{G_{\min}\}$ resulting from Peek's equation 3.

Empirical Equation Representing Another Set of Experimental Data for Sphere Gaps in Oil

Fig. 33.6 of reference 3 shows the results of experiments for 60-cycle spark-over gradients of sphere gaps in commercial oil. The sphere radii were 3.125 and 6.25 cm, and spacings over the range 2.0 cm $\leq x \leq$ 40.0 cm were used. These experiments supplement Peek's results in an interesting way, because the experiments were performed under entirely different circumstances with oil of a different characteristic. The maximum spark-over gradients G were calculated from the spark-over voltages V by means of the same coefficients f (for symmetrical potential distribution, with both spheres insulated from ground) used by Peek.

The spark-over gradients are listed in Table II. For the minimum spark-over gradients $G_{\min}$ a relation having the form of equation 3 is expected

$$\{G_{\min}\}=A+\frac{B}{\sqrt{R}}\qquad(11)$$

Using the values $R_1=6.25$ cm, $G_{1\ \min}=70$ kv$_{\max}$/cm, and $R_2=3.125$ cm, $G_{2\ \min}=87$ kv$_{\max}$/cm, the values of the coefficients A and B are determined so that equation (11) becomes

$$\{G_{\min}\}=29+\frac{102.5}{\sqrt{R}}=29\left(1+\frac{3.53}{\sqrt{R}}\right)\qquad(12)$$

It is seen that the coefficients are not very different from those in equation 3.

Table I. Empirical Equations for Peek's Experimental Data for Sphere Gaps in Oil

$R_1=6.25$ cm; $\{G_1\}=82+90\left(\log\frac{\xi}{1.3}\right)^2$; $[G_1]=73.5+66\left(\log\frac{\xi}{2.5}\right)^2$

ξ	0.0422	0.0813	0.1225	0.1615	0.203	0.285	0.407	0.610
G_1	289	214	180	166	147	130	110	90
$\{G_1\}$	282	212.5	177	155.5	140	121	105	92
$[G_1]$	281.4	219.5	187	166.5	152	132	114.5	98

Spark-over Gradients in Kv$_{\max}$/cm

$R_2=3.12$ cm; $\{G_2\}=98+100\left(\log\frac{\xi}{2.2}\right)^2$; $[G_2]=92.5+93.4\left(\log\frac{\xi}{2.5}\right)^2$

ξ	0.0845	0.163	0.245	0.323	0.407	0.570	0.815	1.220
G_2	288	220	182	166	158	137	122	103
$\{G_2\}$	299.5	226	188.5	167	151.5	132	116.5	104.5
$[G_2]$	295.5	223.5	187	166	150.5	131	115	102

$R_3=1.27$ cm; $\{G_3\}=125+110\left(\log\frac{\xi}{3.5}\right)^2$; $[G_3]=129+\left\{\begin{matrix}146.8\\62.25\end{matrix}\right\}\left(\log\frac{\xi}{2.5}\right)^2$

ξ	0.102	0.208	0.40	0.605	1.00	1.40	2.00	3.00	4.00
G_3	364	310	222	187	158	143	130	131	133
$\{G_3\}$	384	286	222.5	189	157.5	142.5	131.5	125.5	125.5
$[G_3]$	412	300	222	185	152	138.5	130.5	129.5	131.5

$R_4=0.555$ cm; $\{G_4\}=169+170\log\left(\frac{\xi}{4.0}\right)$ [illegible] $[G_4]=180.5+\left\{\begin{matrix}222\\94\end{matrix}\right\}\left(\log\frac{\xi}{2.5}\right)^2$

ξ	0.232	0.476	0.681	1.38	1.82	[illegible].29	3.21	4.57	6.87	9.15
G_4	394	333	295	225	192	[illegible]	171	174	184	192
$\{G_4\}$	429	314	269.5	205	189	179	170.5	169.5	178.5	191
$[G_4]$	418	295.5	251.5	195	184.5	181	182	187	198.5	210

$R_5=0.159$ cm; $\{G_5\}=335+210\left(\log\frac{\xi}{3.0}\right)^2$; $[G_5]=312+\left\{\begin{matrix}414\\175\end{matrix}\right\}\left(\log\frac{\xi}{2.5}\right)^2$

ξ	0.813	1.245	1.66	2.03	2.38	3.20	4.82	6.35	8.00	11.2	16.0	24.0	32.0
G_5	449	360	365	360	364	368	353	358	361	384	416	470	542
$\{G_5\}$	429.5	384	363.5	350.5	345	336	335.5	345.5	357	381	417	469.5	515
$[G_5]$	510.5	350	325	315.5	312	314	326.5	341	356.5	286.5	426.5	480	526.5

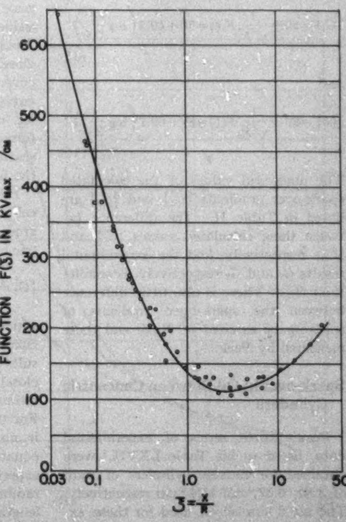


Fig. 2. Encircled points are experimental values of $(G-28.3)\sqrt{R}$. Curve represents function $F(\xi)$, equation 9

Table II. Empirical Equations for Roth's Experimental Data for Sphere Gaps in Oil

$R_1=6.25$ cm $\{G_1\}=70+49.2\left(\log\frac{\xi}{3.0}\right)^2$

ξ	0.4	0.8	1.2	1.6	2.0	2.4	3.2	4.0	4.8	5.6	6.4
G_1	110	88	78	73	71	70	70	70	71	72	73
$\{G_1\}$	107.5	86	78	73.5	71.5	70.5	70	71	72.0	73.5	75.5

$R_2=3.125$ cm $\{G_2\}=87+69.7\left(\log\frac{\xi}{3.0}\right)^2$

ξ	0.8	1.6	2.4	3.2	4.0	4.8	6.4	8.0	9.6	11.2	12.8
G_2	110	92	87	87	88.5	91	96	101.5	107	113	120
$\{G_2\}$	110	92	87.5	87	88	90	94.5	100	105	110	115

ξ	0.4	0.8	1.2	1.6	2.0	2.4	3.2	4.0	4.8	5.6	6.4	8.0	9.6	11.2	12.8
$F_1(\xi)$	202.5	147.5	122.5	112.5	105	102.5	102.5	102.5	105	107.5	110				
$F_2(\xi)$		143		112.5		102.5	102.5	105	109.5		118.5	128	138	148.5	161
$F(\xi)$	196.5	143.0	122.0	111.5	106.5	103.5	102.5	104.5	107.5	111.6	116.0	125.0	134.0	143.0	151.5

In a similar way as for equation 8 two functions $F_1(\xi)$ and $F_2(\xi)$ are now defined by

$$(G_1-29)\sqrt{R_1}=F_1(\xi) \text{ and } (G_2-29)\sqrt{R_2}=F_2(\xi) \quad (13)$$

The numerical values listed in Table II show that the difference between these two functions is relatively small and therefore both can be approximated by means of one function $F(\xi)$

$$F(\xi)=102.5+123\left(\log\frac{\xi}{3.0}\right)^2 \quad (14)$$

Substituting this value in equation 13 the spark-over gradients $\{G_1\}$ and $\{G_2\}$ can be obtained

$$\{G_1\}=29+\frac{1}{\sqrt{R_1}}F(\xi)=70+49.2\left(\log\frac{\xi}{3.0}\right)^2 \quad (15)$$

and

$$\{G_2\}=29+\frac{1}{\sqrt{R_2}}F(\xi)=87+69.7\left(\log\frac{\xi}{3.0}\right)^2 \quad (16)$$

The numerical values of the calculated spark-over gradients $\{G_1\}$ and $\{G_2\}$ are listed in Table II. The difference between these calculated values $\{G_1\}$ and $\{G_2\}$ respectively, and the experimental results G_1 and G_2 respectively, is smaller than 5%. There is also little difference between the spark-over gradients of sphere gaps as cited by Roth and those measured by Peek.

Spark-over in Oil Between Concentric Cylinders

Peek's three series of experimental data, listed in his Table LXVII,[1] were obtained for outside cylinders of radii of 3.81, 6.67, and 11.42 cm respectively. The no. 6 Transil oil used for these experiments is characterized by a dielectric strength of 58 kv_{max} between standard disks 0.5 cm apart. Since the values of the spark-over gradients between spheres shown in Table LXIV refer to no. 6 transil oil with 58.5 kv_{max} spark-over voltage between standard disks 0.5 cm apart, it might be assumed that both series of experiments were performed with oil of similar characteristics. Peek mentioned that the tests with $R=6.67$ cm, and $R=11.72$ cm were made on short cylinders where the field was somewhat disturbed. However, if the spark-over gradients G_1 which occur at the surface of the inner cylinder with the radius r are plotted against log r (a logarithmic scale for r is used to show clearly small values of r), the experimental values for $R=3.81$ cm and $R=6.67$ cm can be closely approximated by the same smooth curve, which is indicated in Fig. 3 as solid curve G. For $R=11.72$ cm only a smaller number of experiments were performed. They show a relatively high scattering and they will not be used for the following consideration.

Peek derived for $R=3.81$ cm an empirical equation for the maximum spark-over gradients (in kv_{max}/cm) between concentric cylinders

$$\{G\}=36\left(1+\frac{1.2}{\sqrt{r}}\right)=36+\frac{43.2}{\sqrt{r}} \quad (17)$$

which is shown in Fig. 3 with the dashed curve $\{G\}$. It fits the experimental results for $R=3.81$ cm and for $R=6.67$ cm closely as long as values of the inner radius r larger than 0.2 cm are considered. For 0.032 cm the experimental value G is about 50% higher as predicted by equation 17; see Table III. From Peek's experiments it seems to follow that the radius R of the outside cylinder and the length L of both concentric cylinders influence only very little the values of the spark-over gradient G.

Analytical Form of the Theory of the Equivalent Volume

Up to now only purely empirical forms for the 60-cycle spark-over gradients in commercial oil have been discussed. Experimental results have been represented by means of simple equations without any attempt to find basic principles. As mentioned before, it does not yet seem possible to use the results obtained recently for spark-over in highly purified liquids and for impulse corona in commercial oil for these problems. Therefore, the only resort at present is to apply the theory of probability for numerical calculations of the spark-over in oil.

Wilson[15] has developed the theory that the primary factor which determines the dielectric strength of commercial oil of controlled quality around electrodes of different geometries is the volume of that oil under stress. Applying considerations which are basically statistical, he concludes that for small volumes of oil, statistical variations are large and unit dielectric strengths are high. For large volumes at wide spacings, lower unit dielectric strengths must result, but statistical reliability and reproducibility are greatly improved. Beyond these qualitative statements Wilson attempts to show that the numerical values of the dielectric strength of any geometrical electrode arrangement may be obtained from test data on a sphere gap with the same equivalent volume. The equivalent volume V_e is enclosed between each electrode and the equigradient surfaces at which the stress is about 90% of the maximum value. For oil of a given characteristic (considering oil quality, humidity,

Table III. Peek's Experimental Values of Spark-over Gradients for Concentric Cylinder (R = 3.81 Cm) Compared with Peek's Empirical Equation and with Values Calculated with Theory of Equivalent Volume

r/cm	0.032	0.238	0.317	0.635	0.794	0.952	1.111	1.270	1.587	1.905	2.540
G/kv$_{max}$/cm	420	127.7	108.1	86.3	85.5	78.1	79.4	77.0	75.1	70.7	62.4
$\{G\}$kv$_{max}$/cm	278	124.5	112.5	89	84.5	80.5	77	74.5	70.2	67.5	63
$[G]$/kv$_{max}$/cm	186	100	101.5	86.5	82.5	79	76.5	74.5	71	68.5	65

dissolved gas, solid and liquid contamination, and including details of the performance of the experiments such as rate of voltage rise, rest time between tests, etc.) there should be one definite function

$$[G]=f(V_e) \tag{18}$$

Peek's empirical equation 3, already used, states that $\{G_{min}\}$ for sphere gaps in oil is a linear function of the reciprocal value of the square root of the sphere radius R. Wilson assumes that the spark-over gradient G_{min} occurs for a gap spacing x_0 approximately equal to $2R$, or $\xi_0=2.0$, and in this case the equivalent volume is found to be approximately

$$V_e=0.4R^3 \text{ or } R=1.35\sqrt[3]{V_e} \tag{19}$$

Since it is assumed that only one definite function $f(V_e)$ determines the spark-over gradient G, there must be, in general, for an oil whose quality is characterized by equation 3

$$[G]=f(V_e)=28.3\left(1+\frac{4}{\sqrt{1.35}\sqrt[3]{V_e}}\right)=28.3+\frac{97.5}{\sqrt[3]{V_e}} \tag{20}$$

This analytical form, not contained in Wilson's paper, yields the same values as shown in his graph 2. It is supposed to hold for oil of a special characteristic which is determined by the numerical values shown in Peek's Table LXIV. This oil will be designated as oil of characteristic 1 in subsequent discussion. In general, one expects for commercial oil of another characteristic (cf. equation 11) an expression

$$[G]=A+\frac{B/\sqrt{1.35}}{\sqrt[3]{V_e}} \tag{21}$$

This should be valid for any electrode arrangement provided that the definition for the equivalent volume already defined can be applied, for instance, for sphere gaps with any gap spacings. It cannot be used for electrodes with sharp edges (see reference 3, page 133, Fig. 33/8), because in this case the gradient at the edge assumes an infinite value.

For symmetrical sphere gaps with small spacings ($\xi \leq 0.1$) Wilson indicates an equation[15] for the equivalent volume V_e (his equation 7), which can be written

$$V_e=0.372\xi^3R^3 \tag{22}$$

For $\xi=2$, the value $V_e=0.4\ R^3$ has already been used. In general, for any ratio $\xi=x/R$ the equivalent volume V_e can be expressed by

$$V_e=S(\xi)R^3 \tag{23}$$

This equation is a similarity law.

If equation 23 is substituted in equation 21 it follows for the spark-over gradient of sphere gaps

$$[G]=A+\frac{B/\sqrt{1.35}}{\sqrt[3]{S(\xi)}\sqrt{R}} \tag{24}$$

For small spacings the special form

$$[G]=A+\frac{B/\sqrt{1.35}}{\sqrt[3]{0.372}\sqrt[3]{\xi}\sqrt{R}} \tag{25}$$

is obtained.

The equivalent volume of concentric cylinders with an inner radius r and the length L is

$$V_e=0.75Lr^2 \text{ (cf. Wilson's equation 8)} \tag{26}$$

Therefore the spark-over gradient (with concentric cylinders) is

$$[G]=A+\frac{B/\sqrt{1.35}}{\sqrt[3]{0.75L}\sqrt[3]{r}} \tag{27}$$

Finally, for parallel, cylindrical disks with the radius R and the spacing s, an equivalent volume $V_c=R^2\pi s$ is obtained, and therefore, (with cylindrical disks)

$$[G]=A+\frac{B/\sqrt{1.35}}{\sqrt[3]{\pi s}\sqrt[3]{R}} \tag{28}$$

It is seen that for the most important cases of electrode geometries the theory of the equivalent volume brought in these equations, which are not contained in Wilson's paper, gives relatively simple laws for the spark-over gradients. These equations contain only two constants, A and B, to characterize the quality of the oil used, the rate of voltage rise, etc. In this way an approximate picture is obtained of possible forms of the equations containing the geometrical parameters which influence the dielectric strength of oil.

Theory of Equivalent Volume Applied to Sphere Gaps in Oil

Wilson's Table II shows calculations of the dielectric strength $[G]$ for spheres of 3.33-cm radius at small spacings. Using equations 23 and 20 the equivalent volumes V_e' and V_e'' and the expected values of the spark-over gradients $[G']$

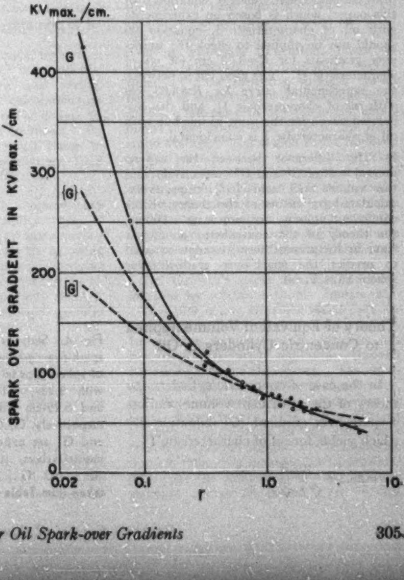


Fig. 3. Sixty-cycle spark-over gradients of concentric cylinders in oil. Points indicate Peek's experimental data for R=3.31 cm, small circles for R=6.7 cm. G is best-fitting curve for these values. {G} is calculated with Peek's empirical equation 7, [G] is calculated with equation 30

Table IV. Equivalent Volumes and Calculated Spark-over Gradients for Spheres with 3.33-, 3.12-, and 6.25-Cm Radius

ξ	2.0	0.946	0.45	0.156	0.075	0.03	0.015
V_e/cm³	14.8	5.8	2.12	0.37	0.077	0.0124	0.0031
[G]	90.5	101	114.5	143	177.5	230	282.5
V_e'/cm³	12.2	4.78	1.75	0.305	0.0635	0.0102	0.00255
[G']	92	102.5	116	147	182	237	290
V_e''/cm³	98	38.4	14.0	2.45	0.51	0.082	0.0205
[G'']	74	81	91	112	137	176	214

and [G''] for spheres with the radii 3.12 and 6.25 cm can be calculated. These values are listed in Table IV. In Fig. 4 the spark-over gradients [G], [G'], and [G''] as predicted by the theory of the equivalent volume, and the experimental values G (taken from Peek's Fig. 181), and G' and G'' (taken from Peek's Table LXIV) are plotted versus log ξ. Fig. 4 shows the following results:

1. Peek's experiments with the spheres with $R=3.33$ cm were not performed with oil of characteristic I. The expression "oil of characteristic I" considers oil quality, humidity, etc., and details of the performance of the experiments such as rate of voltage rise, rest time between tests, etc. Otherwise the experimental curve G for $R=3.33$ cm could not intersect the curves G' for $R=3.12$ cm and G'' for $R=6.25$ cm, but should be located between these two curves, and very close to the first. The oil used by Peek in his experiments with $R=3.33$ cm is designated here as oil of characteristic II, in contrast to the oil of characteristic I used for the experiments listed in Peek's Table LXIV.

2. Since both equation 20 and Wilson's corresponding graph (his Fig. 2), are derived from the numerical values of equation 3 of this paper, which is based on experiments with oil of characteristic I, equation 20 should not be applied to check the spark-over gradients for $R=3.33$ cm for oil of characteristic II. The good check between the experimental curve for $R=3.33$ cm with oil of characteristic II, and the calculated curve for the same radius but with oil of characteristic I is coincidental.

3. The difference between the experimental values G' and G'' respectively and the values [G'] and [G''] respectively, calculated by means of the theory of the equivalent volume, are very large. Hence, the theory of the equivalent volume, at least in its present form, cannot be used to predict the spark-over gradients for sphere gaps in oil.

Theory of Equivalent Volume Applied to Concentric Cylinders in Oil

In the case of concentric cylinders the theory of the equivalent volume predicts a spark-over gradient [G] (equation 27) which yields for oil of characteristic I

$$[G]=28.3+\frac{102.5}{\sqrt[6]{L}\,\sqrt[3]{r}} \qquad (29)$$

This equation compared with Peek's equation 17 results in the following:

1. The spark-over gradient should not depend on the outer radius R, according to both equations.

2. In equation 29 the gradient is a linear function of the reciprocal value of the third root of r, whereas in Peek's equation 17 the reciprocal value of the square root of r is used.

3. Equation 29 shows a weak influence of the length L of the cylinders (reciprocal value of the sixth root); in Peek's equation the length L is not contained explicitly.

In his paper, Wilson discusses the most probable length L used in Peek's experiments with $R=3.81$ cm, because no recorded values of the length could be found at that time. By careful study of details still available he concluded that the length was about 75 cm. With this value equation 29 becomes

$$[G]=28.3+\frac{50}{\sqrt[3]{r}} \qquad (30)$$

Fig. 3 and Table III show the values [G] of the spark-over gradient computed with equation 30. For inner radii $r \geqq 0.25$ cm the approximation is very good. For smaller values of r equation 30 should not be applied.

Major Problems in This Field

As mentioned previously, reliable experimental data for 60-cycle spark-over in commercial oil with different electrode geometries are very scarce. To achieve better results, refinements of the experimental technique are necessary. Even when the greatest care is taken to observe the same testing procedure (rate of voltage rise, etc.) with the same electrode geometries and with oil samples of the same qualities, the results obtained at different places and by different observers show much larger differences than might be expected.[23] It would be desirable to have series of experiments similar to those done by Peek in 1915 carried out on the basis of the improvements in testing techniques with commercial oil which have been established since that time.

Furthermore, it will be necessary not only to record the mean values of spark-over gradients but also to observe carefully the statistical distribution of these measurements. Even though it has been shown here that the application of the "weakest link" theory of probability, as presented by the theory of the equivalent volume, cannot be used successfully in this form, it is desirable to find a general principle, based on the theory of probability, which may lead to simple mathe-

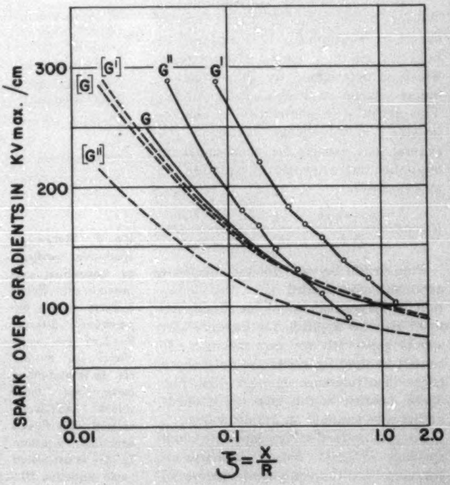


Fig. 4. Sixty-cycle spark-over gradients of sphere gaps in oil with 3.33-, 3.12-, and 6.25-cm radius respectively. G, G′, and G″ are experimental values; [G], [G′], and [G″] are taken from Table IV

matical expressions for the spark-over gradients in commercial oil. Such theoretical and experimental work on spark-over with various gap geometries might result in an improved method of testing the dielectric strength of oil.

Conclusions

1. An empirical equation is derived for Peek's experimental data on sphere gap spark-overs in commercial oil. This equation also represents with good accuracy another series of experiments with sphere gaps in oil of different quality when different numerical values of the coefficients are used.

2. When Peek's empirical equation is applied to several sets of experimental data for predicting the spark-over gradients of oil between concentric cylinders, a good correlation can be obtained for only inner radii r larger than 0.2 cm. The need for further experiments is apparent.

3. It is shown that it is not possible to check the experimental data for sphere gaps and concentric cylinders by calculations based on the theory that the dielectric strength depends on the equivalent volume of oil. However, this theory, when expressed in simple analytical forms and applied to different gap geometries, is shown to be a valuable help in obtaining an approximate picture of possible forms of equations containing the geometrical parameters which influence the dielectric strength of oil.

4. The paper points out the necessity for more and better-controlled experiments with various gap geometries which might result in an improved method of testing the dielectric strength of oil.

References

1. Dielectric Phenomena in High-Voltage Engineering (book), **F. W. Peek, Jr.** McGraw-Hill Book Company, Inc., New York, N. Y., third edition, 1929.

2. Dielectric Phenomena II—Electrical Discharges in Liquids (book), **S. Whitehead.** Ernest Benn, Ltd., London, England, 1928.

3. Hochspannungstechnik (book), **Arnold Roth.** Julius Springer, Vienna, Austria, third edition, 1950, Fig. 33.6, p. 132.

4. Elektrophysik der Isolierstoffe (book), **A. Gemant.** Julius Springer, Berlin, Germany, 1930.

5. Das flussige Dielektrikum (book), **A. Nikuradse.** *Ibid.*, 1934.

6. Dielectric Breakdown of Solids (book), **S. Whitehead.** Clarendon Press, Oxford, England, 1951, p. 3.

7. Some Results on the Electrical Breakdown of Liquids Using Pulse Techniques, **W. D. Edwards.** *Canadian Journal of Physics*, Ottawa, Ontario, Canada, vol. 29, 1951, p. 310.

8. Electrical Conduction and Breakdown in Liquid Dielectrics, **D. W. Goodwin, K. A. MacFadyen.** *Proceedings*, Physical Society, London, England, pt. B, vol. 46, 1953, p. 85.

9. **A. H. Sharbaugh, J. K. Bragg, R. W. Crowe.** *Annual Report*, Conference on Electrical Insulation, National Academy of Sciences, National Research Council, 1953, p. 16.

10. The Dielectric Breakdown of Some Simple Organic Liquids, **B. Salvage.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 98, 1951, p. 227.

11. Reference 9, p. 1.

12. Propagation Mechanism of Impulse Corona and Breakdown in Oil, **T. W. Liao, J. G. Anderson.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 641–48.

13. Reference 2, p. 129.

14. Insulation Variability: A General Equation for Strength Distribution Curves, **M. C. Holmes.** *Journal*, Franklin Institute, Philadelphia, Pa., Feb. 1948.

15. A Fundamental Factor Controlling the Unit Dielectric Strength of Oil, **W. R. Wilson.** *AIEE Transactions*, vol. 72, pt. III, Feb. 1953, pp. 68–74.

16. The Dielectric Strength of Air, **A. Russel.** *The Philosophical Magazine*, London, England, vol. 11, 1906, p. 256.

17. Cf. reference 1, p. 27.

18. Dielectric Phenomena—Electrical Discharges in Gases (book), **S. Whitehead.** Ernest Benn, Ltd., London, England, 1928, diagram XVII, p. 64.

19. Elektrische Durchbruchsfeldstaerke von Gasen (book), **W. D. Schumann.** Julius Springer, Berlin, Germany, 1932.

20. Reference 18, diagram XX, p. 66.

21. Reference 1, equation 29, p. 73.

22. A Note on Standard Calibrations for Sphere-Gaps, **S. Whitehead.** *Journal*, Institution of Electrical Engineers, London, England, vol. 84, 1939, p. 408.

23. Der Einfluss verschiedener Elektrodenformen auf die Durchschlagsspannung von Transformatorenoelen, **L. Maurer.** *Elektrotechnische Zeitschrift*, Wuppertal, Elberfeld, Germany, vol. 74, April 11, 1953.

◆

Discussion

W. R. Wilson (General Electric Company, Philadelphia, Pa.): Dr. Gauster's paper is concerned largely with discussion of my paper[1] which introduced what is called the equivalent-volume or volume-probability theory of the dielectric strength of oil. A second paper,[2] by Wilson, Streater, and Tuohy, describing practical application of this volume theory, was presented concurrently with Dr. Gauster's paper.

Before proceeding with this discussion, the reader is urged to consider carefully the following typical experimental fact. Dielectric strength has been determined for two uniform-field test gaps (shown in the closure of reference 2 of this discussion), each with 0.1-inch metal-to-metal clearance in the same oil at the same time. The plane area of the larger electrodes, and therefore the volume of oil under stress, was 100 times that of the smaller electrodes. In air, the two gaps would have essentially the same dielectric strength. In oil, the small volume gap had an average dielectric strength of 36 kv rms, but the large volume gap with the same metal-to-metal clearance had an average dielectric strength of only 22 kv rms. This large reduction of average dielectric strength is a reproducible result of the probability relationships of the first volume-theory paper[1] and can be estimated from equation 1 of that paper. The standard deviation for the small volume gap was 6.0 kv; and equation 1 yields a predicted value of 21 kv for the large volume gap in comparison with the 22-kv test value.

In the two volume-theory papers, the point of view has been as follows:

1. The unit (or spatial maximum) dielectric strength of commercial oil decreases with increasing volume of oil under stress.
2. It is important to know about how large the effect is, and the direction of its results.
3. As indicated, for example, by certain inconsistent data in Peek,[1] other effects, at present unknown, are also acting. Therefore, one should base calculations on the known performance of a design as similar as possible to the unknown. By so doing, one can come much closer to correct practical design predictions than he can if he ignores this effect.

As is shown in the two volume-theory papers, the volume-probability type of calculations check with surprising accuracy the test dielectric strengths for the following conditions: spheres at moderate spacing (Peek[1]); concentric cylinders (Peek[1]); parallel disk gaps at small spacings (Peek[1]); 3.33-cm radius sphere gap at small spacings (Peek[1]); disk gaps at large spacings (Peek[1]); spherically ended rods at large spacings (Minor[3]); capacitors with different lengths of roll;[2] and numerous practical oil circuit breakers.[2] It is concluded that these excellent checks constitute strong evidence for the existence and importance of the volume-probability effects.

However, as one of the evidences that additional factors are present, reference 1 of the discussion points out, that "even in this source (Peek), careful cross-checking between different test series reveals discrepancies. For example, plotting the data from Fig. 181 and from Table LXIV yields a pair of curves which cross."

The stated area of application of the two volume-theory papers was for the large spacings and large radii of curvature of oil circuit breakers. From reference 2 of this discussion, "reducing clearances and reducing electrode dimensions result in marked increases in statistical deviations. . . . the dielectric strength curves become so steep that for electrodes of small radius, a few per cent of error in the radius might make 50-per-cent error in the withstand voltage. This region is not involved in the problem under study, except that the calculation method and physical picture are supported by the fact that they predict both the instability and the large increases in dielectric strength." The region of Dr. Gauster's comments concerning the volume theory is that of small spacings and small radii.

Dr. Gauster's paper can now be summarized by reference to his figures. Fig. 1 is a plot of the aforementioned data on sphere gaps from Peek's Table LXIV. In addition to plotting the data, he has mathematically fitted equations, through equations 4(A), to the curves. The equations have no stated physical significance, and yield about the same information as curves drawn by eye. Fig. 2 is a different mathematical curve-fit of the same data (through equation 10) and also is applicable only to them. It includes the approximation that the minimums in the dielectric strength curves all lie at $X/R = 2.5$.

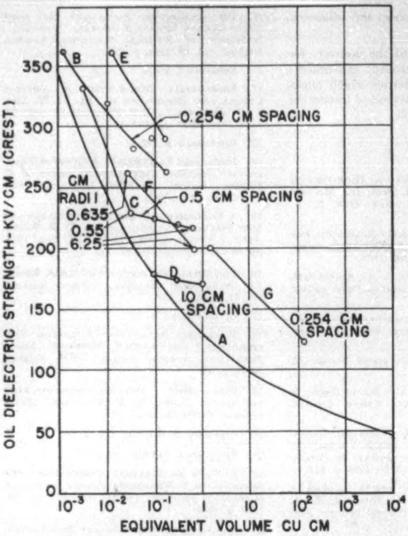


Fig. 5. Curves of dielectric strength versus volume

A—Composite curve from Peek's graphs B, C, D—From Gauster's equations, for sphere gaps at fixed spacings indicated E, F—Modern data for 0.635- and 6.25-cm radius spheres at spacings indicated G—Modern data for plane parallel gaps at 0.254-cm spacing

None of the equations can be applied to cylinders or other more general practical configurations. Depending on the condition of the oil, both the author's equations (when they apply) and the particular volume-theory curve presented will yield results that may be either higher or lower than test data by a factor of more than 2 to 1.

Fig. 3 for concentric cylinders includes the only data (except for corroborative data by Roth) not referred to in the first volume-theory paper. These data consist of the points for the three cylinders with the smallest radii ($R = 0.16$ cm and under). These points were discarded by Peek and by the writer in plotting the same curve.

Fig. 4 is the exact set of curves referred to in the first volume-theory paper: "The data from Fig. 181 and Table LXIV (Peek) yield a pair of curves which cross."

Now, in examining this figure and table which are the basis of Figs. 1, 2, and 4, one sees that Peek presented the table for the purpose of establishing the minimums in the dielectric strength curves of sphere gaps versus spacing (which occur at moderate spacing, or a spacing to radius ratio of, roughly, 2.5). These values, plotted by Peek on his graph 182, check very well with the volume theory (see Fig. 3, reference 1). However, when he discussed dielectric strengths at small spacing, he abandoned Table LXIV and used data with many more points to cover the same range. These he plotted as Fig. 181.

Thus, at low spacings, Peek rejected the tabulated data and selected those of the figure. These (and other) inconsistent results were noted in the first volume-theory paper and it was concluded that "because of the large difference in the dielectric strength of oil under conditions which are superficially identical, the absolute values cannot be applied with safety in engineering calculations for modern oil under any specific set of conditions." Dr. Gauster, however, reverses Peek, rejecting the figure data as incorrect, but then selecting the tabulated data as correct, basing his equations on them. Incidentally, both sets of data agree quite well at moderate spacing. Dr. Gauster then concludes, based on the poor correlation between the volume theory and the low-spacing tabulated data, "Hence the theory of the equivalent volume, at least in its present form, cannot be used to predict the spark-over gradients for sphere gaps in oil." And, "The good check between the experimental curve for $R = 3.33$ cm [low-spacing data selected by Peek] . . . and the [volume theory] calculated curve . . . is coincidental."

This actually lays the groundwork for an interesting example of the value of the volume theory, because it can be shown that either the volume theory is wrong, or Dr. Gauster's equations (based in part on the tabulated data rejected by Peek) are wrong at low spacings. This conclusion results simply from obtaining dielectric strength values from Dr. Gauster's equations for spheres of different radii but constant metal-to-metal spacing, and plotting them against equivalent volumes in Fig. 5. If the spacing-to-radius ratio is small, the field is essentially uniform. However, because the larger spheres will have a larger region under stress, according to the volume theory they should have a lower breakdown voltage. The results as seen in Fig. 5 indicate strong dependence of dielectric strength on equivalent volume at 0.254-cm spacing. The absolute values, of course, are quite different from those of the original paper. However, at 0.5- and 1.0-cm spacings, there are flat regions where dielectric strength does not depend appreciably upon volume, according to Dr. Gauster's equations. If this is factual, the volume theory is wrong.

To check this point, new data were taken. The oil was of course different from that used by Peek. A typical test was as follows: Two sphere gaps, one with large spheres and one with small, were placed in the same oil at the same time. Voltage was raised to breakdown, first on one gap, then on the other, until 80 test values were accumulated. This method of test permits reliable relative data to be obtained. If data are not taken with some such continuous control, the results may be remarkably inconsistent. Table V shows the data, together with the result of volume-theory calculation (see equation 1 of reference 1) of the breakdown gradient for the large spheres, based on the data for the small spheres. Many similar tests were run with similar results. Those reported in this discussion were run at 60-cycles, rapid rise.

According to Dr. Gauster's equations, the relative gradients for these two spheres at 0.5-cm spacing should be almost identical, instead of being different by 30 per cent; see Fig. 5. Consequently, it is concluded that both Gauster's equations and the data Peek rejected in Table LXIV are incorrect at low spacings, and contain inconsistencies. It should perhaps be noted that, although the absolute values of the modern data are well above those of the volume-theory curve based on Peek's data, the slopes based on the pairs of points determined under the same conditions are quite similar.

The volume theory predicts a further interesting corollary. According to it, all curves of sphere gap breakdown voltage versus spacing for spheres of different radii should cross, because at any given very low spacing the smaller spheres will have the higher breakdown voltage, but at relatively high spacing the larger spheres will have sufficiently better uniformity of field that they will have the higher breakdown voltage.

Peek's data of Table LXIV do not show this crossing. It is assumed that there was little interest in the low-spacing values for large spheres, that the series of curves

Table V. Sixty-Cycle Rapid-Rise Tests

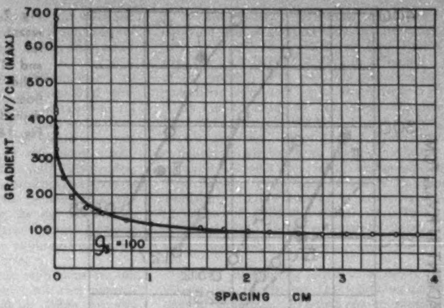
Radius, Cm	Spacing, Cm	Breakdown, Kv Crest	Gradient, Kv per Cm	Standard Deviation	Calculated Gradient	Error, %
0.635	0.254	80.0	359	51		
6.25	0.254	72.9	291		284	3
Test repeated at higher spacing						
0.635	0.508	102	259	48.5		
6.25	0.508	99	199		186	6

was drawn to the origin approximately together (Peek actually shows this in his Fig. 177), and that the points in the table were read from these curves. This was surmised at the time the first volume-theory paper[1] was written, and was the reason for using the figure data rather than the tabulated data at low spacing. This merging of curves produces the flat regions which result when Dr. Gauster's equations are plotted on Fig. 5. Roth's curves and test points (reference 3 of the paper) also show this disinclination to let the curves cross. That the curves actually can cross is shown by the fact that in the aforementioned relative data for 0.254-cm spacing, the smaller spheres had the higher breakdown voltage. Furthermore, in his Fig. 177, Peek shows disk and sphere gap curves which cross in accordance with the volume theory predictions.

Now, concerning concentric cylinders, Dr. Gauster states in conclusion 3, "It is shown that it is not possible to check the experimental data for . . . concentric cylinders by calculations based on the theory that the dielectric strength depends upon the equivalent volume of oil." However, in the body of the paper he states, concerning concentric cylinders, "For inner radii $r \geq 0.25$ cm, the [volume theory] approximation is very good."

In the applications of the volume theory, the radii have been in the order of 5 cm, or in the "very good" range. In fact, the very good range covers all of the data that Peek chose to plot on his graph, and excludes only the three points he discarded. Although several possibilities suggest themselves as to the nature of the unrecognized factor which causes the misfit at low radii, there is one factor which is known to come into play in this general range. When an electrode radius is reduced below a critical value, flashover voltage becomes independent of the radius in a manner similar to a needle gap (see Minor[1]). Therefore, the calculated gradient at flashover becomes a meaningless concept. For example, consider the two largest of the three small cylinders ($R = 0.079$ and 0.159 cm). The two corona starting voltages are 92 and 88 kv, while the two calculated gradients are 251 and 155 kv/cm.

Fig. 6. Peek's Fig. 181 showing increase of strength of oil at small spacings at 60 cycles. (Spheres, R=3.33 cm). By permission of McGraw-Hill Book Company, Inc., reference 1 of the paper



Both Dr. Gauster's paper and this discussion have been concerned primarily with negative considerations. Now it seems appropriate to look at the areas of agreement, because these are so much the greater. Corroboration for the probability effects is found in reference 13 of the paper, which was previously unknown to me. Work by Farmer is described in reference 13 in which he discusses very briefly what he assumes to be an "area effect" for relatively sharp-edged disk gaps in oil. However, he gives no evidence which bears on the question as to whether the variation is not actually with volume. At very small spacings, the volume (or area) effect is so great that breakdown apparently occurs in the low-stress, high-volume, uniform field rather than at the high-stress, low-volume edges. Breakdown voltage decreased with increasing electrode diameter, but no mathematical correlation was attempted. At higher spacings, over about 0.4 inch, the area effect was stated to disappear, obviously because of the increasing relative field concentrations at the sharp edges, in comparison with the average field. As is well known, sharp-edged electrodes (for which spacing is large with respect to edge radius) give radically lower statistical variations than rounded electrodes, so probability effects virtually disappear. Whereas Farmer deals only with average field strengths, unit dielectric strength values at the edges would be many times greater and of totally different form. Further corroboration of the existence of a volume theory was found on page 61 of reference 2 of the paper just as this discussion was being completed. A general though uncertain dependence on volume is postulated, together with speculation that sphere gap curves should cross.

Table VI. Peek's Table LXIV, Dielectric Strength of No. 6 Transil Oil—Spark-over Between Spheres at 60 Cycles*

	Radius of Spheres, Cm										Needles
	0.159		0.555		1.27		3.12		6.25		0
Spacing, Cm	Kv Max	Gradient Max, Kv/Cm	Kv Max	Gradient Max, Kv/Cm	Kv Max	Gradient Max, Kv/Cm	Kv Max	Gradient Max, Kv/Cm	Kv Max	Gradient Max, Kv/Cm	Kv Max
0.129	44.6	449	48.0	394	47.1	364					
0.198	50.5	360									
0.264	59.2	365	73.9	333	73.5	310	74.3	288	74.8	289	
0.322	65.0	360									
0.378	71.1	364	90.0	295							
0.508	81.0	368			99.0	222	105.0	220	107.0	214	
0.650											41.6
0.766	89.9	353	106.0	225	117.0	187	128.0	182	132.0	180	
1.010	97.6	358	112.0	192			146.0	166	159.0	166	
1.270	104.0	361	116.0	176	157.0	158	171.0	158	177.0	147	57.0
1.780	111.0	384	131.0	171	165.0	143	203.0	137	214.0	130	
2.540	124.0	416	149.0	174	185.0	130	240.0	122	245.0	110	84.0
3.810	145.0	470	172.0	184	206.0	131	266.0	103	280.0	90	108.0
5.080	168.0	542	191.0	192	231.0	133					124.0
7.620											166.0
10.150											203.0

* Oil between disks 0.5 cm apart tested 58.5 kv max 25 C. By permission of McGraw-Hill Book Company, Inc., reference 1 of the paper.

Judging from Dr. Gauster's oral presentation, some additional areas of agreement are as follows: probability effects are important and real; Peek's data, or data of similar scope, should be taken again with modern controls and reported; it is not certain that the probability effects now attributed solely to volume do not actually depend upon both volume and surface area; the volume theory gives a first approximate answer to what is taking place, but at present there is no known method of calculation which will give accurate results over wide extrapolations. I hope that Dr. Gauster will continue his activities and that he will separate some of the additional variables that certainly seem to exist.

In applying the volume-probability principles in their present form, if one has sufficient data to establish performance of one dielectric structure, he can calculate the performance of another fairly similar structure with good engineering accuracy. This is verified by the low calculation errors for plane and spherical gaps in this discussion, and for the ±5% accuracy realized in the practical work with oil circuit breakers.[2]

REFERENCES

1. A FUNDAMENTAL FACTOR CONTROLLING THE UNIT DIELECTRIC STRENGTH OF OIL, W. R. Wilson. *AIEE Transactions*, vol. 72, pt. III, Feb. 1953, pp. 68–74.

2. APPLICATION OF VOLUME THEORY OF DIELECTRIC STRENGTH TO OIL CIRCUIT BREAKERS, W. R. Wilson, A. L. Streater, E. J. Tuohy. *AIEE Transactions*, vol. 74, pt. III, 1955 (*Paper 55-219*).

W. F. Gauster: Since both Mr. Wilson and I have based our work on Peek's Table LXIV and Fig. 181, it is worth while to reproduce these sources here, as Table VI and Fig. 6. Peek's own comments

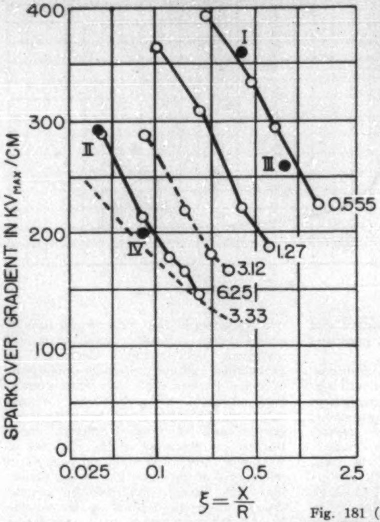


Fig. 7. Points I and III represent Wilson's experiments for R=0.635 cm, points II and IV for R=6.25 cm. Solid lines and dashed line from Peek's Table LXIV, dashed curve from Peek's Fig. 181. Radii shown in cm

(reference 1 of the paper) are as follows: "Sparkover voltages and gradients are given for various spheres at various spacings in Table LXIV. The characteristic of the curves between gradient and spacing are the same as those for air (see Fig. 42) as shown in Fig. 181." And "For the small spacings, the apparent rupturing gradient is very high, just as in the case of air. Strengths as high as 700 kv/cm have been reached." Peek's Fig. 181 (Fig. 6) shows the graph of the spark-over gradients versus spacings; however, small values of spacings can be read with some exactness only if they are not much smaller than approximately $x=0.1$ cm. This spacing corresponds to a spark-over gradient of about 250 kv$_{max}$/cm. For gradients higher than 250 kv$_{max}$/cm and spacings smaller than 0.1 cm this graph cannot be used. Peek's captions of Table LXIV (Table VI) and Fig. 181 (Fig. 6) indicate the different goals which he had in mind when he showed the different experimental data. The captions are: "Table LXIV Dielectric Strength of No. 6 Transil Oil—Sparkover Between Spheres at 60 Cycles" with the remark: "Oil between disks 0.5 cm apart tested 58.5 kv max., 25 deg. C" and, on the other hand: "Fig. 181. Showing increase of strength of oil at small spacings at 60 cycles. (Spheres, $R=3.33$ cm)." No special oil quality, and no table of numerical values of the spark-over gradient are indicated in the latter case.

Peek's obvious intention was to list in Table LXIV a set of consistent numerical values of experimental data for the dielectric strength of a special oil quality at a certain temperature, whereas the purpose of Fig. 181 was to show in a qualitative way only the typical performance at very small spacings. The difference of the two oil qualities is very obvious; oil as used for Fig. 181 ($R=3.33$ cm) has, as indicated, the dielectric strength of about 250 kv$_{max}$/cm for a spacing x approximately equal to 0.1 cm, whereas in Table LXIV ($R=3.12$ cm) as spark-over gradient of 288 kv$_{max}$/cm occurs for $x=0.264$ cm.

Nevertheless there must be no doubt of the "peaceful coexistence" of these very different sets of Peek's experimental data. Such great differences can be easily observed when fresh oil is used for one series of experiments, whereas the other experiments are performed with oil which has already been used for some time, especially when it has been kept in an open tank. Also the details of the performance of the experiments, such as rate of voltage rise, rest time between tests etc., have a great influence. These facts are clearly expressed in my paper when I introduced the designations "oil of characteristic I" and "oil of characteristic II" for Table LXIV and Fig. 181 respectively. This is also in accordance with Mr. Wilson's remark (reference 15 of my paper) that the discrepancy between the table and the figure values may be explained "perhaps because of different rates of voltage rise."

There is no doubt that all experimental data concerning dielectric strength of oil have to be used with caution, especially when details in the performance of the experiments are not indicated and when no detailed statistical evaluation of the results is shown. Nevertheless up to now Peek's Table LXIV (showing the 60-cycle spark-overs of sphere gaps in oil) from 1915 seems to be the best set of experimental data of this kind available up to now. As shown in Fig. 1 of the paper all the values of the spark-over gradient plotted against $\log x/R = \log \xi$ seem to be sufficiently consistent with each other. On the other hand, the different deviations of the points from the best-fitting smooth curves show that none of Peek's numerical values were obtained by extrapolation only. In Table LXIV, values which were not actually observed have been omitted. There are also some indications of a phenomenon, predicted with the use of his volume theory by Mr. Wilson in his discussion, namely, that breakdown voltages versus spacings for spheres of different radii should cross at very low spacings. As shown in Peek's table at $x=0.129$ cm, the breakdown voltage for $R=1.27$ cm is somewhat smaller than for $R=0.555$ cm. The same effect may exist at $x=0.264$ cm for $R=0.555$ cm, and $R=1.27$ cm respectively. However, the uncertainty of these experimental data is too high to make any definite statement. This phenomenon probably depends on the rate of voltage rise.

Considering all these facts, I cannot agree with Mr. Wilson's assumptions about Peek's Table LXIV and Fig. 181 as developed in his discussion. His hypothesis can be summarized as follows: For larger spheres Peek's Table LXIV should be correct only as far as minimum spark-over gradients are concerned. For smaller spacings of larger spheres the tabulated values should be rejected, since they were obtained by Peek by means of incorrect extrapolations. Finally the spark-over gradients shown in Fig. 181 should supplement the unrejected tabulated values in a consistent way.

The check of Mr. Wilson's volume theory for sphere gaps with small spacings would be correct only if the hypothesis mentioned in the foregoing could be accepted. Since this is not the case, my view is maintained that Mr. Wilson's check of his volume theory for small sphere spacings is not valid.

This opinion is strengthened by the results of Mr. Wilson's new experiments with sphere gaps. In Fig. 7 the spark-over gradients (in kv$_{max}$/cm) are plotted against $\log x/R = \log \xi$ similar to Figs. 1, 2, and 4. Points I and III represent Mr. Wilson's experiments for $R=0.635$ cm, points II and IV for $R=6.25$ cm. The solid lines represent Peek's tabulated values for $R=0.555$, 1.27, and 6.25 cm respectively; the dashed line represents Peek's tabulated values for $R=3.12$ cm. The dashed curve is from Peek's Fig. 181 for $R=3.33$ cm. Mr. Wilson's new experimental values for $R=6.25$ cm are very close to Peek's tabulated values for this radius, especially point II with the smaller spacing. Since Peek did not make experiments with a spacing of $x=0.635$ cm, his tabulated values for 0.555 cm and 1.27 cm are shown in Fig. 7. In case of a perfect coincidence of Mr. Wilson's values with Peek's tabulated data, points I and III ($R=0.635$ cm) should be located between these two curves, very close to that for $x=0.555$ cm. In fact, Wilson's value III has such a position. Point I is somewhat higher than expected from Peek's tabulated values, but all the differences are relatively very small and may be easily explained by different oil quality, etc., including the fact that Wilson made his experiments with rapid voltage rise. On the other hand, the curve from Fig. 181 for $R=3.33$ cm is even lower than Wilson's experimental value for $R=6.35$ cm. These details give additional proof that the hypothesis of Mr. Wilson about Peek's tabulated values and Fig. 181 does not agree with the actual facts.

Fig. 5 shows that for $x=0.5$ cm Mr. Wilson's experimental values are between 9% higher and 14% lower than those calculated with my empirical equation for the oil characteristic as used for Peek's Table LXIV. At a spacing $x=0.254$ cm the difference is between 10% and 18%. These differences are greater than those obtained when Peek's experimental data are directly used. This follows from Table I of the paper, where the differences between the values from the empirical equations and Peek's tabulated values are listed. As mentioned in this connection in the paper, the accuracy of empirical equation 9 is sufficient as far as experimental data are available up to now. As soon as better experimental data are obtained, empirical equation 4 should be used and "the next goal should be to find simple empirical forms for the coefficients $N(R_k)$ and $\xi_0(R_k)$ similar to that in equation 3 for $G_{\min}$."

Whereas the magnitude of the differences between Mr. Wilson's new experiments and Peek's tabulated values, are from 0 to 10%, and from 0 to 18% when using my experimental equation, the differences between Mr. Wilson's new experimental values I, II, III, and IV and his unit dielectric curve derived by means of the volume theory from Peek's minimum gradients (which are accepted by Wilson for this oil quality) are 56, 77, 26, and 46% respectively.

To obtain a new, complete check of the volume theory as presented by Mr. Wilson, it would be necessary to establish, for the oil quality used and the special way of performance of the experiments involved, the curve of unit dielectric strength from experiments with spheres of large spacings. It would not be sufficient to show that lines corresponding to E and F in Fig. 5 coincide with each other. They also should coincide with the unit dielectric curve derived from the minimum spark-over gradients. Mr. Wilson's new experiments do not satisfy these requirements.

When the validity of the volume theory is checked, great care must be taken. For instance, a check of Mr. Wilson's volume theory for the case of the oil breakdown strength for "isolated" spheres (large spacings) is not conclusive. Mr. Wilson's theory is based on Peek's empirical equation

$$G_{\min}=28.3\left(1+\frac{4}{\sqrt{R}}\right)kv_{\max}/\text{cm}$$

which holds for the constant part of the curve and therefore Mr. Wilson's curve for the dielectric strength of oil between isolated spheres is only another mathematical expression for Peek's empirical law.

In his discussion, Mr. Wilson refers to experiments with two uniform-field test gaps, each with 0.1-inch metal-to-metal clearance in the same oil at the same time. The plane area of the large electrodes, and therefore the volume of oil under stress, was 100 times that of the smaller electrodes. I consider this not an appropriate experimental arrangement to demonstrate the validity of the volume theory. The large difference of the dielectric strength in these two cases could also be explained easily by means of a surface effect.

Concerning concentric cylinders in oil, Peek writes (reference 1 of the paper): "It is much more difficult [than for air] to detect the starting point of corona and, unless the conductors are very small or far apart (s/r large), corona does not appear before spark-over. For instance, in the Table LXVII when the outer cylinder is 3.81-cm radius and the inner 0.0127-cm radius ($R/r=300$), the corona and spark-over voltages are practically the same. . . . The spark-over voltages and corona voltages up to fairly high ratios R/r are, therefore, the same and may be used in determining the strength of oil."

Peek disregarded the values $r=0.032$, 0.159, and 0.079 cm, not because in these cases corona occurs before spark-over, but because his empirical equation

$$\{G\}=36\left(1+\frac{1.2}{\sqrt{r}}\right)=36+\frac{43.2}{\sqrt{r}} \qquad (31)$$

gives results much too low for these values of r. As shown in Fig. 3 of the paper, the values derived by means of the volume theory are still lower.

With small inside radii r the spark-over gradients as listed in Peek's Table LXVII show a very strong dependence on r, whereas the curve representing the breakdown voltage V_s becomes flat. However, if the volume theory yields a spark-over gradient of about 145 $\text{kv}_{\max}$/cm instead of 251 $\text{kv}_{\max}$/cm, the predicted breakdown voltage is instead of 88 kv (crest) only

$$88\,\frac{145}{251}=51 \text{ kv (crest)}$$

which is about 42% too low.

Peek's empirical equation for concentric cylinders seems to be less satisfying than his empirical equation for the minimum spark-over gradients of spheres (equation 3 of the paper). Whereas this latter equation yields good results in the range from $r=0.159$ to $r=12.5$ cm (see Peek's Fig. 182 and Table LXVIII in reference 1 of the paper), Peek uses his equation for concentric cylinders in a range from $r=0.238$ cm to $r=2.540$ cm only. See his Figs. 180(A) and 180(B). For $r=5$ cm, Peek's empirical equation for concentric cylinders yields too high spark-over gradients and Wilson's value is still higher (about 30% higher than the empirical results available). The range of a satisfying coincidence of Wilson's values and Peek's empirical data is approximately between 0.25 cm$\leq r \leq$2.5 cm only; compare Table III.

Before concluding this discussion, I should like to express the following more general thoughts. The first step in physical research is to undertake experiments. In the field of breakdown in oil the need for further experiments is apparent. The second step usually is to establish empirical equations, without any immediate physical significance. For various groups of experiments, different empirical equations must be established and there will be no direct connection between these equations. The third step is to understand which physical factors are responsible for the empirical results. The ultimate goal is attained if some basic equations of physical significance can be established which yield, as special cases, equations which can be used instead of the empirical equations.

Sometimes it is a relatively easy matter to obtain the desired results by application of a general principle without the necessity to consider the physical processes in complete detail. Examples are the principle of the conservation of energy and many variation-principles, which are often applied to problems in physics.

Mr. Wilson's theory of the equivalent volume is an attempt to understand and to predict numerical results for breakdown phenomena in oil, with the use of a general principle, without trying to follow the physical processes in detail. Mr. Wilson applied in a very clear and simple way the concept of spatial probability to breakdown problems. He carefully studied the problem of whether his theory in its actual form is sufficient to check experimental results or whether other factors should be considered. There is hope that the consideration of the volume effect in an appropriate form, combined with that of other probability effects (for instance surface effects) will finally lead to a more complete theory, which will yield the desired results.

Gas and Moisture Equilibriums in Transformer Oil

R. B. KAUFMAN
MEMBER AIEE

E. J. SHIMANSKI
NONMEMBER AIEE

K. W. MACFADYEN
NONMEMBER AIEE

OIL IS the major electrical insulation component of large high-voltage transformers. It must withstand, free of corona, the severe short-time electrical stresses imposed during factory test or abnormal service conditions, as well as the less severe normal-service stresses of long duration. It follows that best insulation design demands precise knowledge not only of the stresses imposed on the oil during test and operation, but also a sound understanding of the oil's ability to withstand successfully the stresses under a wide variety of adverse conditions.

Pure oil, free of all contaminating influences, has very high intrinsic strength, many times that practically and economically attainable in "working" transformer oil. The oil in a transformer always contains some water in solution; certain amounts of dissolved gases are unavoidable. As the transformer is operated through its normal daily and seasonal load cycles, both the water and gas content of the oil are subject to rather wide change. For optimum application of insulation, the designer must know, among other things, how water and gas content of oil are influenced by temperature, pressure and humidity; and how these two contaminants, in turn, influence the electrical strength of the oil.

Unlike air, which has such well-defined and consistent electrical strength behavior that it is frequently used for measuring voltage, the electrical strength of oil is erratic and capricious. Its apparent strength, even when left undisturbed, may change appreciably from one hour to the next, and it is not uncommon for a series of consecutive breakdown determinations made over a 15-minute interval to exhibit such broad statistical dispersion as to leave little confidence in the results. Oil strength may be influenced not only by dissolved water and gas, but by ever-present solid particles, as well as other extraneous factors which include temperature, pressure, and the nature of the electric field.

Many dollars' worth of time and effort have been expended on studies of the electrical strength of oil. Some of the factors that affect oil strength were identified qualitatively several decades ago and they include, of course, dissolved water and gas. It seems safe to say, however, that no really clean-cut, quantitative, and usable conclusions have been reached. On the contrary, the literature on these subjects is replete with contradiction.

There are three reasons for this, all involving experimental difficulties which can be stated briefly as: 1. recognition and control of all variables, 2. quantitative determination of the variables; and 3. interactions among the variables.

For several years a small group of chemists, physicists, and engineers in the General Electric Power Transformer Engineering and Laboratories Department have been working to alleviate these difficulties. Progress has been made in all three areas, and several successful well-controlled experiments have been completed which shed considerable light on oil breakdown behavior and the influences of extraneous factors. This success is due partly to the development of methods of sampling and handling oil without changing it, of precisely determining water and gas content, and of finding basic solubility relationships and equilibrium criteria for gas, water, and oil. These methods and findings constitute the subject matter of this paper in which thorough treatment is given to experimental methods employed to provide clearer understanding and validation of data presented.

Specifically, these highly transient variables have been brought under closer laboratory observation than ever in the past.[1–3] In its standardization and establishment of basic moisture and gas solubility relationships in ranges of common practical interest, this work has laid the foundation for continuing advanced study of these variables directed to a more complete understanding of their practical significance in the design, construction, and maintenance of liquid-filled electrical apparatus. It should be noted that all the experimental data on oil herein reported have been obtained on transformer oil from one particular source.

Analytical

ANALYSIS FOR WATER

Since 1935, when Karl Fischer used an iodine reagent for determining moisture in sulfur dioxide, an ever-increasing number of applications have been described until at present the Fischer reagent is the one most widely used for quantitative moisture determination. Mitchell and Smith's book, "Aquametry,"[4] comprises an excellent an comprehensive survey of the field of water analysis with particular emphasis on the Karl Fischer reagent.

Although the technique utilized for many years comprised a visual determination of color change from chromate-yellow to a red-brown iodine color, this has been found inadequate for the accuracy and precision required in any basic study involving minute quantities of water. For this reason it was decided to evaluate several of the existing electrometric procedures for end-point determination. A circuit suggested by G. R. Cornish[5] was modified and found superior to many of the more elaborate circuits described in recent years, particularly from the standpoint of simplicity and reliability. This modified circuit is shown in Fig. 1, and has been successfully employed in conjunction with a moisture-measuring apparatus in all phases of the program.

Apparatus

The apparatus is shown in Fig. 2, and consists of an all-glass titration chamber into which a burette tip, platinum electrodes, and a Teflon-covered stirring bar are fitted. An automatic burette and reservoir are utilized to titrate accurately with iodine reagent into the sealed system. The electrodes are connected to the detecting circuit by copper leads. The microammeter dial comprises the indicating instrument for end-point detection. A magnetic stirrer is used to agitate the solution during the determination.

Paper 55-66, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing December 3, 1954.

R. B. KAUFMAN, E. J. SHIMANSKI, and K. W. MACFADYEN are with the General Electric Company, Pittsfield, Mass.

The authors are indebted to J. R. Meador and E. L. Raab, General Electric Company, and to C. D. Cook, now at the University of Vermont, for their valuable suggestions and encouragement in carrying out this work. They would also like to thank T. W. Hall for his help in modification and construction of the detection instrument for moisture measurement.

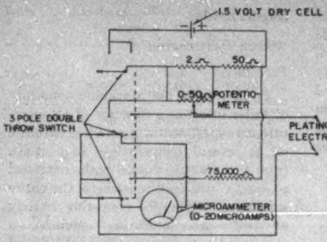


Fig. 1 (left). Circuit diagram for electrometric water method

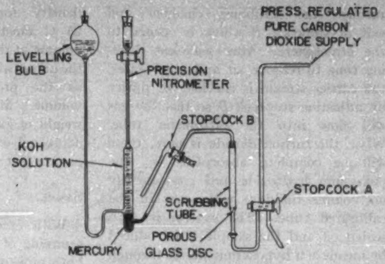


Fig. 3 (right). Apparatus for dissolved gas content of liquids

Procedure

A suitable solvent is introduced into the titration chamber and the end point approached by addition of iodine reagent to neutralize the moisture present in the solvent. While water is present to react with the iodine-forming hydriodic acid, the anode remains depolarized and the cathode polarized, allowing no flow of current. When all the water has been reacted, the addition of iodine depolarizes the cathode, resulting in a current flow. The end point is reached when this flow of current results in a steady reading on the microammeter. A sample (15 to 25 grams) of oil or other liquid of unknown moisture content is added, and the system resealed. If moisture is present, it will be indicated by a decrease in the current flow and a movement of the microammeter needle toward zero. The iodine reagent is then added from the calibrated burette until a steady reading on the instrument is again obtained. The amount of reagent in milliliters (ml) is recorded as well as its predetermined moisture equivalence. This information is then used in the following equation to give moisture content in parts per million

$$\frac{\text{ml Fischer reagent}\times\text{moisture equivalence (grams per ml)}\times 10^6}{\text{Sample weight (grams)}}=\text{parts per million water}$$

The method has been found reliable and the precision calculated to be within ±2 parts per million. Of particular importance is the fact that little experience is necessary for an operator to obtain accurate and reproducible results.

ANALYSIS FOR AIR AND OTHER GASES

In the past 10 years there has been increasing emphasis upon measurement and control of dissolved gases in various liquids. It has always been recognized that this factor was significant in certain systems such as the lubricating, hydraulic, and fuel systems of high-altitude aircraft where extreme pressure and temperature changes coupled with too large a gas content could cause considerable gassing. A comparatively recent development, however, is an increasing concern over this factor in the electrical industry, at least where liquid dielectrics in apparatus of more complex design are concerned.

A number of analytical methods for measurement of dissolved air in insulating liquids have been in use for some time. The majority of these methods, where accuracy and precision are desired, are built around very elaborate high-vacuum apparatus. In most cases the determination involves either absorbing measured quantities of air into a completely degassed oil or measuring the quantity of air evolved when an aerated oil is exposed to a vacuum system.[6] The difficulties encountered in maintaining an absolutely airtight system, calibrating the entire volume of the system, and ensuring equilibrium, make these methods very tedious, expensive, and not as reliable as the method which is used by the authors and hereafter described.

Paralleling the work on moisture already described, and utilizing the same sampling techniques, a method has been developed for determining air and many other gases in liquids such as transformer oil. Fairly inexpensive equipment is involved and the method allows rapid determinations using a small amount (10 ml.) of sample. Results obtained can be expressed either on a volume per-cent basis or on a weight per weight basis and a precision of ±0.2 parts per million by weight is attainable. The operating principle involves the use of a carrier gas which displaces the gas to be determined and is itself subsequently absorbed chemically, allowing volumetric measurement of the displaced gas.

Apparatus

The apparatus is shown in Fig. 3 and consists of an all-glass system which is sealed and into which is piped a supply of pure carbon dioxide. A scrubbing chamber fitted with a porous glass disk for dispersion of the carrier gas is the means by which the rapid displacement takes place. A factory-calibrated nitrometer, commonly used for the microdetermination of nitrogen evolved in organic combustion methods and containing a strong solution of potassium hydroxide, is utilized to collect and measure the gas carried over by the carbon dioxide. The carrier gas in turn reacts chemically with the hydroxide solution and is completely reabsorbed from the gas stream.

Procedure

Stopcock *A* in Fig. 3 is adjusted so that the carbon dioxide is allowed to flow

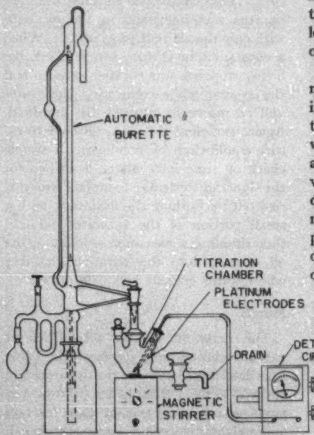


Fig. 2. Apparatus for electrometric water determination

through the scrubbing chamber and exit at stopcock *B* which is opened to the atmosphere. After sufficient flushing time to remove air and other gases, the carrier stream is checked for purity by adjusting stopcock *B* so that the gas will flow into the nitrometer tube. When the carbon dioxide is pure, there will be complete absorption by the potassium hydroxide and consequently no volume build-up will occur in the calibrated tube. The system is then sealed off and the sample is introduced by means of a hypodermic needle through a rubber stopper at the top of the scrubbing tube. The carrier gas is then flowed through the oil causing displacement of the dissolved gas by saturating the oil with the highly soluble carbon dioxide. This displaced gas is collected in the calibrated tube and its volume read by means of a levelling bulb which equalizes the pressure inside the sealed tube with the atmosphere.

The temperature, barometric pressure, volume of gas collected, and volume or weight of oil used are recorded, and the amount of dissolved gas calculated according to the needs of the particular investigation, i.e., either as a weight or volume relationship. Correction factors are supplied with each calibrated burette and should be included in the over-all calculation. The initial portion of the calculation involves expressing the collected gas as a volume at some predesignated standard condition. Conditions chosen in this investigation were 760 millimeters (mm) of mercury for pressure and 25 degrees centigrade (C) for temperature. For this expression the calculation becomes

$$\frac{Va\times Pa\times Ts}{Ps\times Ta}=Vs$$

where

Va = actual volume of gas collected in cubic centimeters
Pa = atmospheric pressure in mm mercury, barometer
Ps = standard reference pressure in mm mercury, 760
Ta = temperature, degrees Kelvin, at which Va was measured
Ts = standard reference temperature, degrees Kelvin, 298
Vs = volume of gas corrected to standard reference conditions

This gives the volume of gas based on a standard reference and allows a realistic comparison of results. If the corrected gas volume obtained is divided by the oil sample volume at 25 C and multiplied by 100, the result is the gas content in volume per cent.

For conversion to weight units the density for the particular gas at 25 C must be known and when the volume at that temperature is calculated, the actual weight of gas is obtained simply as the product of the density and volume. Multiplication of this gram weight of gas by 10^6 and division by the actual oil weight in grams will give the gas content in parts per million.

SAMPLING

With exceptions noted in procedure, sampling of all systems under study for both moisture and gas content measurements was carried out by means of a standard glass medical syringe (10 to 30 ml) equipped with suitable valve attachments, needles, and adapters to meet particular needs. This technique has allowed accurate oil sampling from various systems under equilibrium, both pressurized and evacuated at relatively remote locations, and has allowed convenient transport and storage of these samples in preparation for analysis with no compositional alteration. Where necessary, samples were maintained under exact physical conditions existing in a system in question. In sample transfer to the syringe, especially in liquid systems highly saturated with gas, care was taken to avoid separation of gas during transfer. After sampling was completed, if such separation occurred in subsequent handling, separated gas trapped in the syringe was accurately accounted for in final analysis. Syringe transparency has an advantage in that ready examination of the sample can be made at any handling stage to observe oil condition and avoid error. In cases in which light effects were potentially significant, care was taken to shield the syringe properly. In sampling and analysis for moisture content, it was desirable to allow the internal surfaces of the syringe to come as close to equilibrium as possible with the moisture in the liquid under study. This was accomplished by conditioning the syringe prior to sampling in an atmosphere of proper humidity or, more conveniently, by repeated rinsing of the syringe with the liquid in question. Generally speaking, it is important to recognize the previous history of the samplers before use, particularly when repeated use is involved in studies covering wide moisture variations. It is significant that these same techniques are applicable not only to accurate moisture and gas analysis in oil, but also to critical laboratory study of related transient oil quality factors such as dielectric strength, resistivity, and power factor.

Experimental

WATER SATURATION

Procedure

To determine the maximum solubility of water in oil as a function of temperature, an experiment was set up to saturate oil at several temperatures in a range from 12 to 80 C. An oil bath controlled to ±0.1 C was used to immerse the entire saturation apparatus consisting mainly of two water-saturating columns, a glass wool entrainment filter and an oil reservoir, all in series. Each vessel in this train was fitted with a porous glass dispersion disk through which air was passed at a steady rate. Sampling was accomplished by means of a medical syringe which was preheated or cooled in the same bath so as to minimize temperature change in transfer to the titration chamber.

The time required to insure equilibrium was established by sampling after several widely different time intervals. In all cases, the maximum water content was reached before 8 hours of air flow at the prescribed flow rate. Time intervals of 24, 32, 48, and 72 hours were checked and no increase in water content observed. A 24-hour minimum period was utilized for establishing the equilibrium conditions at various temperatures for the dynamic system described. An independent check for maximum water content was made on a nonflowing system to ensure that the results were not a function of the method of saturating. This all-glass apparatus consisted of an oil and a water reservoir connected through a vented glass manifold and provided with an oil drainage cock. For this static system, the apparatus was immersed in the oil bath with only the air vent protruding. When a constant temperature was reached, the 2-way stopcock was rotated so as to seal the system while connecting the water and oil reservoirs through the manifold. Again, sufficient time was allowed to insure equilibrium by conducting a simple check on time intervals as described for the dynamic system. Sampling was carried out by flushing the draincock with a small portion of the saturated oil and then draining a measured amount of the oil directly into the titration chamber where water content was determined.

Results

The maximum water solubilities at various temperatures are shown in Table I, as determined by both the dynamic and static systems. Results are expressed in parts per million by weight, mole fraction and weight per cent(%) of water in oil.

Vol.
Pt. 1
1955

Table I. Maximum Water Solubilities in Transformer Oil

Temperature, C±0.1C	Parts per Million	Mole Fraction†	Weight, %
12	30	0.00046	0.0030
20	44	0.00068	0.0044
30	67	0.00104	0.0067
40	101	0.00156	0.0101
40.5*	103	0.00160	0.0103
50*	149	0.00231	0.0149
60	213	0.00331	0.0213
60*	215	0.00333	0.0215
80*	411	0.00632	0.0411

* These results determined under static conditions.

† Mole fraction is a quotient obtained by dividing the number of moles if water present by the number of moles of oil and water combined.

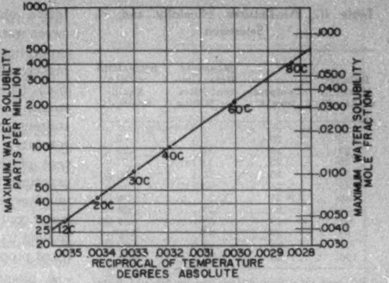


Fig. 5. Water content of transformer oil at saturation as a log function of temperature

These results are plotted in Fig. 4 on rectangular co-ordinate paper and in Fig. 5 on semilog paper showing a linear relationship between log of the solubility and the reciprocal of the absolute temperature.

Discussion

As previously stated, the primary objective of this systematic investigation of moisture is its application to broader study of insulating liquid behavior. Limited consideration is given to theoretical explanation for the observed relationships since this has been the subject of previous exhaustive and authoritative study.[7–9] The quantitative behavior of moisture in oil as observed in this work, although exhibiting some deviation from strict vapor pressure considerations, is generally in accordance with classic laws governing behavior of dilute solutions (i.e., Raoult's and Henry's laws). Although originating in a completely independent investigation, the findings in these experiments agree with those of Black, Joris, and Taylor,[7] demonstrating similar moisture solubility relationships for various pure liquid hydrocarbons. As is indicated in reference 7, such deviations as were encountered can be expected for specific systems of the type in question, i.e., in dilute solutions in which the components in solution exhibit wide differences in internal pressures, polarity, molecular structure, and degree of association in the liquid state.

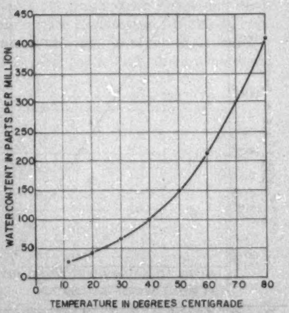


Fig. 4. Water content of transformer oil at saturation as a function of temperature

The work confirmed the fact that as the chemical constitution of hydrocarbon-type liquids change, water solubility in these liquids may also change. However, the characteristic principles cited for transformer oil widely apply to similar organic liquids in establishing moisture saturation-temperature curves and in defining moisture equilibrium as a function of humidity. In this general study, characteristic conformity has been shown not only for conventional transformer oil, a highly complex mixture of hydrocarbons, but also for related liquids of higher or lower aromatic, paraffinic, or unsaturates content, as well as of other differences in molecular structure or composition. This is illustrated in the comparison of the moisture saturation curves for conventional transformer oil with an oil of predominant aromatic composition, and their comparison, in turn, with that of benzene and to published saturation data[10] for cyclohexane, as shown in Fig. 6. It will be noted that all exhibit a linear relationship between the log of solubility and the reciprocal of the absolute temperature. Slopes of the saturation curves vary because of intrinsic compositional differences in the liquids. Characteristic variations of the same type have been observed on transformer oil itself because of differences in natural origin and refining or impurities originating in oxidative or thermal deterioration of a particular oil with use.

Relative Water Saturation

Procedure

The question of water-oil equilibriums for varying air relative humidities was also investigated and the % moisture saturation of oil in equilibrium with moist air was found to agree exactly with the % relative humidity of the air when both are at the same temperature. To express this in another way, it might be said that when oil and moist air are in equilibrium at some given temperature, the ratio of water present in air to the maximum that air will contain at that temperature, is exactly equal to the ratio of water in oil to the maximum water that oil will contain at that temperature. It was important that this phenomenon be studied since there are substances in which water does not absorb according to this rule. Paper and pure cellulose are among such substances in which a nonlinear relationship exists between water absorption and air-relative humidity.

The relative humidity of the air for these experiments was established by use of both static and dynamic methods. In certain cases the oil was placed in a sealed container in which the % of humidity was varied by means of saturated salt solutions. In other instances, a flowing air

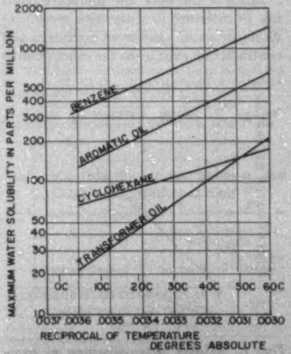


Fig. 6. Water content at saturation for several liquids as a log function of temperature

ol. 74
. 1
955

Table II. Air-Relative Humidity and % Saturation

Relative Humidity of Air, %	Temperature, C	Measured Water Content Parts per Million	Saturation* of Oil at Equilibrium, %
0	25	0	0
23	26	14	23
23	25	12	22
36	26	20	36
45	26	25	45
50	25	27	50
54	25	29	53
57	26	33	59
69	25	38	70
69	26	38	68
100	26	56	100

* Per-cent saturation values are determined from a knowledge of the 100% or saturated water content for the temperatures given. At 25 C this value is 54 parts per million and at 26 C it is 56 parts per million by reference to Fig. 5.

stream was passed through either a glycerine-water solution or a saturated salt-water solution and then dispersed through the oil. Excellent agreement was achieved between these two methods, indicating, as in the case of the 100% saturation data, that the equilibrium is not dependent upon the method used to attain it.

Results

The results are shown in Table II and plotted in Fig. 7. Although very little information is available on relative saturation at elevated temperatures, it is assumed that this linear relationship between air humidity and % water saturation in oil holds throughout the temperature range of practical interest. A family of curves relating water content of oil with air relative humidity over a practical range of temperatures is shown in Fig. 8.

Discussion

Much of the work and literature relating to moisture in dielectrics recognizes its importance. However, confusion and contradiction exist as to the exact quantitative relationships and correlations between water content and various electrical properties. Much of this is due to preoccupation with attempted correlations based on absolute water content without reference to what a given water content represents in terms of relative saturation. This excludes a novel concept of the authors, well substantiated experimentally, that the state of relative moisture saturation, and not absolute water content per se, is a critical and controlling factor with respect to many electrical properties.

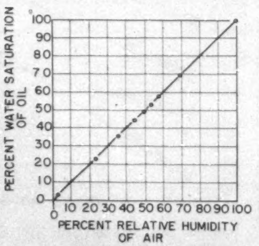


Fig. 7. Relative water saturation of transformer oil as a function of air humidity at equilibrium

A practical illustration of the significance of the relative moisture picture is given by reference to parallel work relating specific electrical properties to relative water saturation. An article on a continuous-flow moisture detector[11] deals in detail with resistance, power factor, and capacitance measurements as a function of and as a means for detecting the water content of a flowing oil system. A typical picture of the relationship between the insulation resistance and the relative water saturation over a fairly wide temperature range is given in Fig. 9. These results, determined by the authors in the course of the fundamental studies on moisture equilibriums and coupled with other findings presented in this paper, provide a foundation for relating many other dielectric phenomena to the relative amounts of water present in a liquid system. If these fundamental relationships are properly recognized and utilized, a more profitable approach can be made to quantitative interpretation of dielectric phenomena.

AIR SATURATION

Procedure

Similar experimentation was carried out to determine maximum air solubilities in transformer oil at various temperatures and pressures. A constant temperature bath and a simple apparatus consisting of a glass vessel with an air vent and containing about 500 cubic centimeters of oil was used. Samples were drawn into a syringe which was preheated for determinations at elevated temperature and then transferred to the analytical apparatus shown in Fig. 3. As in the case of the water saturation study, several time intervals were checked until it was certain that equilibrium had been reached. Temperatures of from 25 C to 120 C were studied. Air saturations of this static oil system, as well as one through which air was passed by means of a porous glass disk, were determined and found to agree, providing that any free air bubbles were allowed to escape.

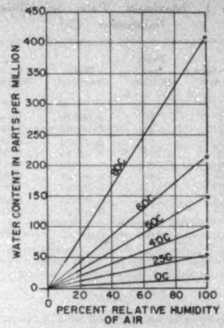


Fig. 8. Water content of transformer oil in equilibrium with moist air for several temperatures

Results

The maximum solubilities of air in transformer oil at atmospheric pressure for temperatures of practical interest are given in Table III. The data, plotted in Fig. 10, show that, at fixed pressure, the solubility of air is directly proportional to the temperature at which it was saturated.

Although a volume % expression is a rather convenient means for picturing the gas condition of a liquid insulator, and is useful for various engineering calculations, it can be misleading if inadequately defined. As one means of expressing gas solubility in the present report, the fol-

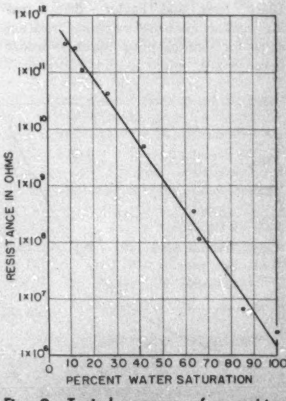


Fig. 9. Typical response of a moisture device—insulation resistance as a function of per-cent water saturation of oil

Table III. Maximum Air Solubility in Transformer Oil for Several Temperatures

Temperature, C	Pressure,* Mm Mercury	Air in Solution at Equilibrium	
		Volume %	Parts per Million
25	736	10.00	135.3
40	736	10.20	137.9
60	736	10.57	142.8
75	736	10.81	146.2
120	736	11.48	155.2

* The oils were saturated at the prevailing barometric pressure which ranged from 727 to 745 mm, and then corrected to 736 mm by reference to Fig. 11.

lowing conditions were chosen for reference: 25 C for oil temperatue; 25 C, 760 mm for gas temperature and pressure. By such definition, therefore, gas solubility as volume per cent is calculated in accordance with the following expression

$$\frac{\text{Gas volume at 25 C, 760 mm}\times 100}{\text{Oil volume at 25 C}} = \text{volume \%}$$

Standardized use and consistent adherence to such a point of reference is highly desirable. In comparison to these defined conditions, selection of different reference conditions of temperature and pressure obviously give rise to apparent different volume relationships between gas and oil which are unreal in the sense that they rest in the method of calculation only. To minimize the possibility of ambiguity, another expression for gas content is suggested. This involves the conversion of gas content of oil to a weight per weight relationship as expressed in parts per million. These units provide an advantage in that they represent a clear-cut measure of solubility independent of volume change. To say that 1,000,000 grams of oil at 75 C contains 146 grams of air in solution needs no further clarification. If this is compared with the figure for 25 C, which is 135 parts per million, there is no doubt about the increase of solubility with temperature.

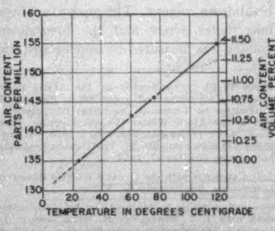


Fig. 10. Air content of transformer oil as a function of temperature

The effect of pressure change on the air content of oil was determined in a separate experiment in which degassed oil was allowed to come to equilibrium with air at pressure levels ranging between 1 mm and 1,000 mm of mercury. This establishes the direct relationship between air content and pressure and also provides a means of correction for barometric change while reaching equilibrium. This curve is in Fig. 11 and can be assumed to hold for pressures up to several atmospheres. Similar linear relationships have been demonstrated for other gases.

In addition to air solubility, a brief study was made of the solution in oil of the individual major components of air, nitrogen, and oxygen. The maximum solubilities of each were determined at 25 C when present as the only dissolved gas component in a previously degassed oil sample. This showed oxygen to be more highly soluble than nitrogen in transformer oil, and led to the investigation of their solubilities as components of air in equilibrium with oil. Although nitrogen comprises 78% by volume of the atmosphere and oxygen 21%, it was found that in oil only 70% of the gas was nitrogen while 30% was oxygen. These data are shown in Table IV. The oxygen enrichment in oil solution is due to the greater solubility of oxygen in oil as also shown in Table IV. It should be noted too that, on a weight basis, oxygen constitutes 40% of air dissolved in transformer oil.

Conclusions

1. Improved analytical sampling techniques and apparatus have been developed and usefully applied to measurement, control, and fundamental study of moisture and gas in transformer oil and related liquids.

2. Fundamental temperature and humidity relationships governing moisture solubility in transformer oil have been established and found in agreement with basic laws governing dilute solution. Data show that the log of the solubility of water in oil and similar organic liquids is directly proportional to the reciprocal of the absolute temperature and that it increases with temperature. This saturation plot provides a useful means of accurately extrapolating limited data to obtain saturation data over the full practical range of interest.

3. Saturation characteristics of these liquids, while conforming to fundamental solubility relationships, show variation attributable to intrinsic differences in individual liquid properties.

4. Relative moisture saturation of transformer oil and various similar liquids has been found to be linearly related to the relative humidity of the atmosphere with which equilibrium was established.

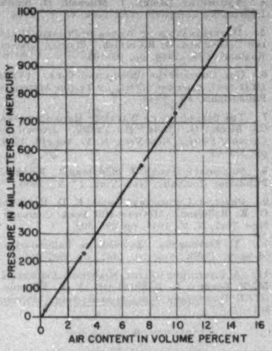


Fig. 11. Air content of transformer oil as a function of pressure

5. Relative moisture saturation, rather than absolute moisture content per se, is suggested as a controlling factor in quantitative correlation of water and many electrical properties of dielectrics.

6. The solubility of air and its major components in transformer oil have been defined for various temperatures and pressures of practical interest.

7. It is shown that air solubility in transformer oil increases slightly with temperature between the limits of 25 and 120 C, and that air undergoes a compositional change in oil solution because of inherent differences in solubility of its major components.

8. The air content of oil at saturation varies linearly with the pressure at which equilibrium is established.

References

1. WATER SOLUTION IN HIGH-VOLTAGE DIELECTRIC LIQUIDS, F. M. Clark. *AIEE Transactions (Electrical Engineering)*, vol. 59, Aug. 1940, pp. 433–41.

2. THE ROLE OF DISSOLVED GASES IN INSULATING OILS, F. M. Clark. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 215, no. 1, pp. 39–67.

3. LIQUID DIELECTRICS (book), A. Gemant. John Wiley & Sons, Inc., New York, N. Y., 1933.

Table IV. Solubility of Air, Nitrogen, and Oxygen in Transformer Oil*

Dissolved Gas	Gas in Solution at Equilibrium	
	Volume %	Parts per Million
Air	10.00	135
Nitrogen only	8.56	112
Oxygen only	15.92	237
Nitrogen as air component	6.98	91
Oxygen as air component	3.02	44

* Equilibrium established at 736 mm mercury and 25 C.

4. Aquametry (book), **J. Mitchell, D. Smith.** Interscience Company, New York, N. Y., 1948.

5. Determination of Water in Plastic Molding Compounds, **G. R.Cornish.** *Plastics*, London, England, Feb. 1946, pp. 99–103.

6. Gas Content of Insulating Oils. *ASTM D831-1948*, American Society for Testing Materials, Philadelphia, Pa.; 1948.

7. The Solubility of Water in Hydrocarbons, **C. Black, G. Joris, H. Taylor.** *Journal of Chemical Physics*, New York, N. Y., vol. 16, 1948, pp. 537–43.

8. Solubility (book), **J. Hildebrand.** Reinhold Publishing Company, New York, N. Y., 1936.

9. Physical Chemistry (book), **E. D. Eastman, G. K. Rollefson.** McGraw-Hill Book Company, New York, N. Y., 1947, pp. 295–302.

10. **T. Berkengeim.** Zavodskaya Laboratoriya, Moscow, USSR, vol. 10, 1941, pp. 592–94.

11. A Continuous-Flow Moisture Detector, **M. F. Beavers, E. J. Shimanski, E. F. Timpane.** *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 361–67).

◆

Discussion

Frank C. Doble (Doble Engineering Company, Belmont, Mass.): We are pleased that the authors understand the concept "that the state of relative moisture saturation, and not absolute water content per se, is a critical and controlling factor with respect to many electrical properties." This concept led me to present before the National Research Council Insulation Conference on November 2, 1939, a paper entitled "Low Loss and High Loss Water in Oil." The terms "low loss" and "high loss water" were suggested for the purpose of describing the phenomena upon which this concept is based.

One of the oils studied by the authors, identified as an aromatic oil, has markedly different water saturation characteristics from that of the transformer oil. Since natural aromatics play an important part in determining the characteristics of transformer oil, we should like to ask whether the aromatic oil is of the transformer oil type and what percentage of aromatics it contained; also if the authors have studied the effect of relatively small changes in aromatic content of transformer oil such as occur when a distillate is subjected to different amounts of refinement.

R. B. Kaufman, E. J. Shimanski, and K. W. MacFadyen: The authors are gratified with Mr. Doble's concurrence on the significance of relative moisture saturation in liquid dielectrics. His implication, however, that this concept was clearly recognized as early as 1939 is rather surprising in view of the industrywide need for, and persistent efforts toward, improved standardization and refinement of methods for more accurately determining the moisture content and dielectric liquids. As is pointed out in the paper, the clear delineation of moisture and its effects in the diminutive solubility ranges and homogeneous systems with which we are concerned required the most careful techniques of sampling and analysis. The many contradictions in published literature and lack of attention to the exact parameters which must be used in definition of moisture solubility and its related effects in dielectrics are difficult to reconcile with Mr. Doble's viewpoint. We have been unable to obtain a copy of his unpublished 1939 paper, to compare his concept with ours of low-loss and high-loss water in oil.

The aromatic oil referred to by the authors was a synthetic blend of essentially complete aromatic composition which was cited simply as a further illustration of a relationship between composition and moisture solubility in hydrocarbon-type liquids. It was not a transformer-type oil in the classical sense to which Mr. Doble refers. Although we have not studied exhaustively the moisture saturation characteristics of every commercial transformer oil, we do know that severe alteration in the aromatics content of those oils with which we are most familiar significantly affects their moisture solubility characteristics. Further, we recognize that this is not the only compositional variation influencing the moisture saturation characteristics of natural oils of this type.

Some Electrical and Thermal Characteristics of Helium and Sulfur-Hexafluoride Mixtures

J. D. COBINE
MEMBER AIEE

Synopsis: The results are presented of a study of some of the electrical and thermal properties of mixtures of helium and sulfur-hexafluoride which are compared with the pure gases: air, hydrogen, helium, and sulfur-hexafluoride. It is shown that the addition of relatively small amounts of sulfur-hexafluoride to helium will result in an electric strength greater than that of air and of hydrogen. The thermal properties of such mixtures are nearly as good as hydrogen. The mixture is attractive as a cooling and insulating medium for both static and rotating machines. Such a mixture is very safe whereas hydrogen presents an explosive hazard. Permeation losses through metal containers are much less than for hydrogen.

HYDROGEN is often used as a gaseous cooling medium in electric machines[1,2,3] where high heat-transfer and low windage power losses are desired. This gas has some disadvantages, among which may be noted the following. Its electrical breakdown strength is about 65 per cent that of air over a considerable range of pressures. There is a certain amount of hazard involved caused by the possibility of air mixing with the hydrogen, either inside the machine or in pockets outside. The permeation rate of hydrogen from metal containers is greater than for other gases.[4]

Helium has nearly as good thermal properties as hydrogen; its loss through metals is very low compared to hydrogen, but its electric strength is very low. It is now well known that the presence of small amounts of electronegative gases tends to raise the breakdown voltage of gases.[5] However, there appears to have been no study of the possibility of developing a gaseous mixture having both high electric strength and good thermal properties. The following summarizes the results of a study of the electrical and thermal characteristics of mixtures of helium and the electronegative gas sulfur-hexafluoride.

Experimental Results

All electrical and thermal measurements were made in a 9-inch bell jar. The base plate and internal arrangement are shown in Fig. 1. All spark breakdown voltages were for 60-cycle-per-second power and were measured for 1-inch-diameter hemispheres of stainless steel which had a 1/8-inch radius of curvature at the back edges. To reduce the erosion of the surfaces of the spheres to a minimum, a thyratron-controlled relay was used to remove the voltage upon the flow of breakdown current. The gap was set by a micrometer screw and the breakdown voltage was measured by an electrostatic

---

Paper 55-10, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted August 25, 1954; made available for printing October 27, 1954.

J. D. Cobine is with the General Electric Research Laboratory, Schenectady, N. Y.

The author wishes to acknowledge helpful discussions with J. K. Bragg, M. H. Hebb, and F. J. Norton. The thyratron-controlled relay described in the paper was developed by S. Reynolds of this laboratory.

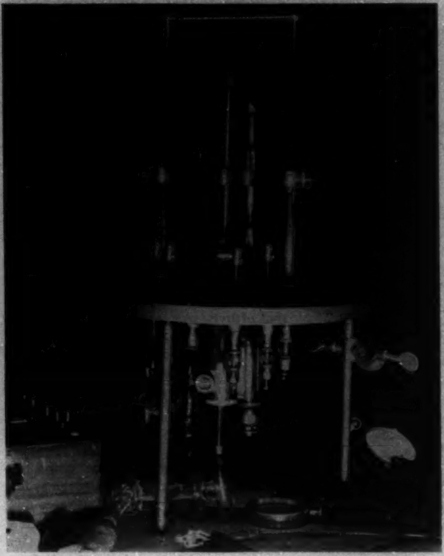


Fig. 1 (left). Experimental apparatus

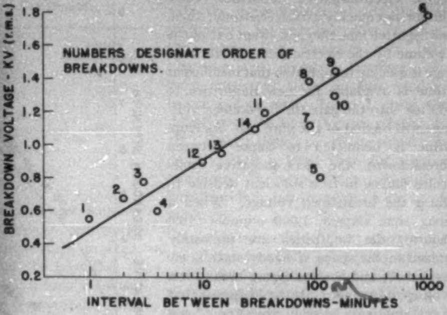


Fig. 2 (above). Effect of interval between breakdown on helium with small amount of SF_6 impurity (74.6 centimeters Hg)

voltmeter (calibrated in rms voltage). The gases, air, hydrogen, and helium were admitted to the system through a dry-ice–acetone trap. The sulfur-hexafluoride was chemically pure and showed no trace of impurities when examined by mass spectrograph, while the helium used was of welding grade. The voltage was increased slowly so that at least 15 seconds were required to reach breakdown. An interval of at least 1 minute elapsed between successive breakdowns. The first spark breakdown data were taken with air so that the equipment could be tested by comparing results with published data.

While taking exploratory data, it was found that once sulfur-hexafluoride had been used in the bell jar, it was impossible to obtain consistent data with helium even after repeated flushings with pure helium. Once contaminated, the breakdown voltage obtained with helium depended markedly on the time permitted to elapse between breakdowns. The data presented in Fig. 2 show that there is quite a systematic dependence of breakdown voltage of helium on the time permitted to elapse between breakdowns. Point 6, at 1,000 minutes, represents the upper limit of breakdown voltage and point 1, taken 1 minute aftere a preceding breakdown, is approximately the lower limit. This lower limit is equal to the value obtained for helium before any sulfur-hexafluoride was used in the system. The effect was certainly caused by residual amounts of sulfur-hexafluoride in the system. The flushing alone would have reduced the gaseous proportion of sulfur-hexafluoride to an extremely small value; however, large surface areas are present upon which many layers of gaseous impurities are adsorbed. Thus, no realistic estimate of the amount of sulfur-hexafluoride present could be made.

The following hypothesis appears to offer a reasonable explanation. When a breakdown takes place in a gas containing

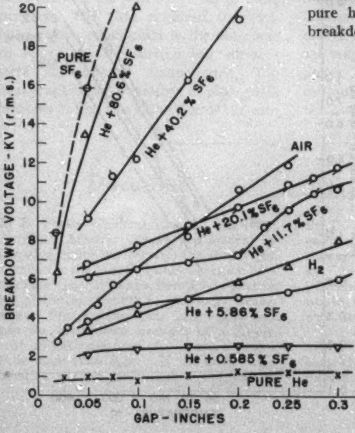


Fig. 3 (left). Spark breakdown voltages for air, He, H_2, SF_6, and mixtures of He+SF_6 for 1-inch-diameter hemispheres. (74.6 centimeters Hg, 29.5 degrees centigrade, 60-cycle voltage)

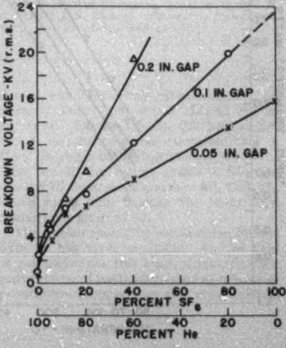


Fig. 4 (right). Spark breakdown voltage for SF_6+He mixtures. Parameter = gap between 1-inch-diameter hemispheres (74.6 centimeters Hg, 29.5 degrees centigrade) some electronegative impurities, these impurities quickly attach electrons; then, as negative ions they are swept out of the volume by the electric field. If the voltage is quickly reapplied so that insufficient time is available for new impurities to diffuse into the gap, the breakdown voltage will be that of the pure gas. As more time is permitted to elapse between breakdowns, the electronegative molecules diffuse in to a sufficient density to raise the breakdown voltage. When a long time elapses, 1,000 minutes (16.6 hours), the impurities are uniformly mixed in the space, a steady state is established, and the maximum breakdown voltage is observed. It would appear that a systematic variation in spark breakdown such as this represents a good way to detect the presence of very small amounts of electronegative gases mixed in other gases. The procedure followed in taking the final data consisted in filling the system to nearly atmospheric pressure and pumping to 200 microns for 5 cycles. In at least 3 cycles, a low-pressure glow discharge of about 10 minutes' duration was established in the new gas. This procedure removed residual impurities from all surfaces to a suitable minimum value.

The 60-cycles-per-second spark breakdown voltages for air, hydrogen, helium, sulfur-hexafluoride, and mixtures of the last two are shown in Fig. 3. The volumetric proportions of the two gases are indicated as percentages. A total pressure of 74.6 centimeters Hg was used throughout the study. The values plotted represent the average of ten breakdowns. The rms deviations from the average value were generally from 1 to 5 per cent, being greatest for the shortest gaps. The breakdown curve for air is slightly lower than published values, probably cause by the fact that the electrodes were not true spheres. However, the ratio of the breakdown voltage of hydrogen to that of air is equal to the published values, namely, about 65 per cent. These curves show the marked effect of even small amounts of sulfur-hexafluoride in increasing the breakdown strength of helium. The data of Fig. 3 are replotted in Fig. 4 to bring out directly the effect of the amount of sulfur-hexafluoride mixed with the helium. As the gap length approaches the radius of the electrodes, the electric field becomes more distorted. For this reason, comparative values are given for only the shorter gap lengths when the field is more nearly uniform. The values of breakdown voltage for the helium–sulfur-hexafluoride mixtures are shown in Fig. 5 relative to that of hydrogen and of air for two gap lengths. The marked gain in electric strength of helium with small additions of sulfur-hexafluoride over both air and hydrogen is brought out quite well by this figure. This gain is especially great for small spacings, as the 0.05-inch-gap data show.

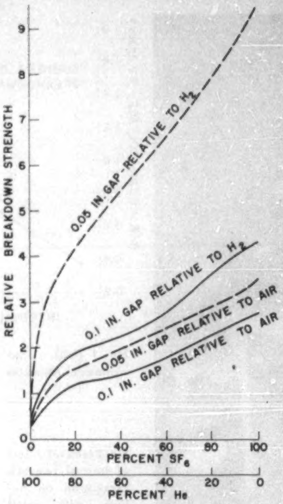


Fig. 5. Relative breakdown strength of mixtures of $He+SF_6$

The total heat loss from structures immersed in a gas is dependent on the geometry and orientation of the surfaces and is usually quite complex. It was felt that a determination of the heat transfer for the several gases relative to air from measurements made in the same simple structure would be instructive. The relative heat-transfer data thus obtained would be useful in estimating the heat losses from other structures where the loss rate in air is known. For the present study, the term "heat transfer" is taken to mean the electric power input to a helix

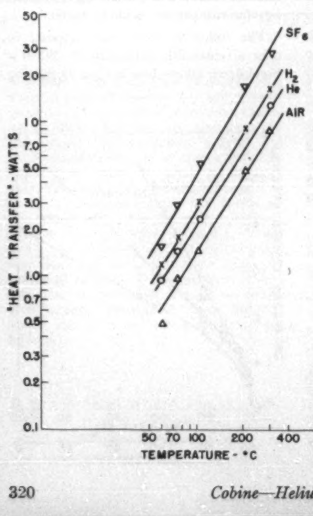


Fig. 6 (left). Heat transfer by free convection and radiation for pure gases, (power in watts required to maintain temperature T at center of wire helix, ambient = 26 degrees centigrade, pressure = 74.6 centimeters Hg)

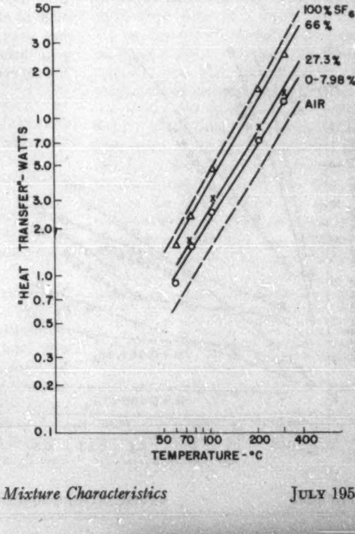


Fig. 7 (right). Heat transfer for mixtures of SF_6+He (ambient = 26 degrees centigrade, pressure = 74.6 centimeters Hg)

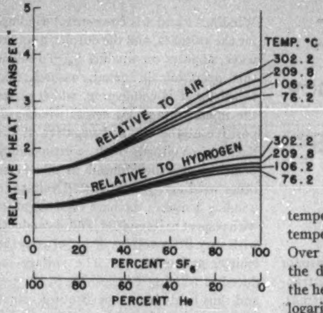


Fig. 8. Relative heat transfer of He+SF$_6$ mixtures

of resistance wire at a given wire temperature. This includes the combined action of free convection in the gas and radiation as well as a small conduction loss through the end leads.

A helix of 0.015-inch-diameter nichrome wire wound on a 5/16-inch mandrel, as shown in Fig. 1, had a platinum-rhodium thermocouple spot-welded at the center. The length of the helix was 15/16 inch, with a pitch of 0.0445 inch. The cold resistance, including that of the contacts, was 5.7 ohms. The length of wire is quite long so that the heat loss of the leads is small relative to the total loss. The power input necessary to maintain a constant temperature of the wire at the center of the helix will be called the heat transfer.

As shown in Fig. 1, the helix is in a large bell jar with a heavy stainless-steel plate; consequently, the heat loss is all to these surfaces. The time required to reach temperature equilibrium in the helix was so short that the surrounding surfaces were at constant temperature. The heat transfer for a number of pure gases is shown in Fig. 6 as a function of the wire temperature, at the constant ambient temperature of 26 degrees centigrade. Over the range of temperature involved, the data indicate that the logarithm of the heat transfer is a linear function of the logarithm of the wire temperature. The values were quite reproducible except at the lowest temperature, 60.4 degrees centigrade, where the measurements were somewhat erratic.

Table I. Gas Densities

Gas	Density Relative to Air
H_2	0.0695
He	0.138
Air	1.0
SF_6	5.03

The heat-transfer characteristics of mixtures of helium and sulfur-hexafluoride are shown in Fig. 7. The relations for air and pure sulfur-hexafluoride are included for comparison. In Fig. 8, the heat-transfer of the He+SF$_6$ mixture is shown relative to air and hydrogen as functions of the per-cent sulfur-hexafluoride. The latter figure shows that the heat transfer of the mixture is always greater than that for air; while it is never less than 80 per cent of that for hydrogen it is better than hydrogen for percentages of sulfur-hexafluoride greater than about 30 per cent.

Discussion

Some of the factors involved in a rotating device, such as an alternator or a gyro, are different from those involved in a static device, such as a transformer or ballast reactor. In a large high-speed machine, the power loss caused by windage may be as much as one-half the total power loss. A cooling and insulating medium should be such as to minimize this loss as much as possible, consistent with the desired electrical and thermal properties. In turbulent flow of gases, the power loss increases with both viscosity and density. However, the most important factor in this windage loss is gas density. Table I gives the densities of the several gases relative to air. Since drag is proportional to density, it is evident that, for roating machines, the percentage of sulfur-hexafluoride in helium must be kept to a minimum.

In addition to the inherent safety in the use of the chemically inert gases, helium, and sulfur-hexafluoride, the mixture has the advantage of being adaptable to the use of a simple halogen leak detector.[6] Permeation loss of both gases through metals is immeasurably low; however loss of helium by permeation does take place through glass, silica, rubbers, and plastics.[4] The combination of an inert gas such as helium with an inert electronegative gas such as sulfur-hexafluoride to produce a gas whose electrical and thermal properties can be tailored to a particular application should have many applications.

References

1. Standard Handbook for Electrical Engineers (book), **A. E. Knowlton**, editor. McGraw-Hill Book Company, New York, N. Y., 8th ed., sect. 7 to 138, 1949, p. 689.

2. Hydrogen as a Cooling Medium for Electrical Machinery, **Edgar Knowlton, Chester W. Rice, E. H. Freiburghouse.** *AIEE Journal*, vol. 44, 1925, pp. 724–34.

3. Diffusion of Gases Through Solids, **C. J. Fechheimer.** *Electric Journal*, East Pittsburgh, Pa., vol. 26, 1929, p. 405.

4. Helium Diffusion Through Glass, **F. J. Norton**, *General Electric Review*, Schenectady, N. Y., Sept. 1952, p. 28. Also, *Journal*, American Ceramic Society, Columbus, Ohio, vol. 36, 1953, p. 90.

5. The Effects of Pressure on the Positive Point-to-Plane Discharge in N_2, O_2, CO_2, SO_2, SF_6, CCl_2F_2, A, He, and H_2, **H. C. Pollock, F. S. Cooper.** *Physical Review*, New York, N. Y., vol. 56, 1939, p. 170.

6. Electronics Simulates Sense of Smell, **W. C. White, J. S. Hickey.** *Electronics*, New York, N. Y., vol. 21, March 1948, p. 100.

◆

Discussion

R. E. Plump (General Electric Company, Pittsfield, Mass.): There are several interesting points for comment in this paper. The first is the fact that once again we have evidence that SF$_6$ has a beneficial effect on the dielectric strength of relatively weaker gases. Fig. 4 shows that in the 0.05-inch gap the breakdown strength of a 20-per-cent SF$_6$–80 per-cent He mixture was a little more than 6 kv. If the linear combination of each gas is made (i.e., assuming direct proportionality) such as, by the equation $kv = x(DS_{SF_6}) + (1-x)\ (DS_{He})$ in which x is the fraction of SF$_6$ present, and the DS's are the dielectric strengths of each gas alone in the gap, then the strength of the mixture should have been only about 4 kv. In other words, the mixture had a 50-per-cent higher dielectric strength than would have been obtained proportionally from each component in the mixture.

The fact that heat transfer with SF$_6$ is higher than with either H$_2$ or He may be somewhat surprising when SF$_6$ is known to have a slower rate of diffusion. However, it also has interatomic vibrations (i.e., an infrared absorption) which H$_2$ and He do not have. Further, the specific heat of SF$_6$ is only one-eighth to one-twentieth that of He or H$_2$ respectively, but in a given volume there are 36 times more grams of SF$_6$ and He and 73 times more grams of SF$_6$ than H$_2$, which more than balance the lower heat content per gram of SF$_6$.

The author should be commended for the data he has made available. In the field of gas dielectrics this kind of investigation proves to be of particular value as industrial possibilities become more and more feasible.

A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage

G. H. ROYER
ASSOCIATE MEMBER AIEE

THE combination of transistors and transformers constructed of square hysteresis loop material has produced much useful circuitry outside of the realm of magnetic amplifiers per se. In this paper a highly efficient and extremely precise converter for transforming direct to alternating current, yielding a square-wave output whose frequency is proportional to the d-c input voltage, will be described. From among the many uses of this converter, the following applications will be discussed: 1. a transducer for telemetering systems, and 2. an integrator which has found application as an analogue to digital converter.

Transistor Switches

The use of junction transistors as controlled switches has been described previously,[1,2] but a few of the main points will be reviewed here. Fig. 1(A) shows the collector characteristics of a PNP transistor in the common emitter connection, as illustrated in Fig. 1(B). For the load line shown, two operating points are of interest: cutoff, point *F*, where the transistor sustains almost the entire supply voltage *E* with very little leakage current and conduction, point *S*, where a large enough base drive is applied to saturate the transistor and reduce the voltage drop V_{ce} to its lowest possible value. Cutoff occurs when the base voltage is slightly positive giving a small positive current in the base. Typical values of cutoff current for small 200-megawatt *2N74* fused junction transistors are from 5 to 50 microamperes at a supply voltage of 45 volts and a base voltage of +0.5 to 5 volts. In the conducting state voltage drops of 0.1 to 0.5 volt are obtained with 50 to 150 milliamperes' load current; however, the amount of base current required to saturate the transistor depends greatly on the gain of the transistor and the amount of load current. The effective transistor impedance can be changed from over a megohm at cutoff to a few ohms at saturation by control of the base voltage and current; this can be compared to the operation of a relay where the emitter-collector circuit is analogous to the contacts and the base-emitter circuit is analogous to the energizing coil. For the purpose of this paper, the above information can be summarized by two simple rules. The PNP transistor switch will conduct if the base is negative with respect to either emitter or collector and will block if the base is positive with respect to both emitter and collector.

Basic Circuit of Converter

Two junction transistors used as controlled switches and a transformer with a square hysteresis loop core material form the basic circuit of the voltage-to-frequency converter illustrated in Fig. 2. A typical ϕ-i curve for such a transformer constructed of Hipernik V is also illustrated in Fig. 3. The two transistors alternately conduct and switch the input voltage V_{in} across windings 2 and 3, causing the flux in the core to be cycled between positive and negative saturation.

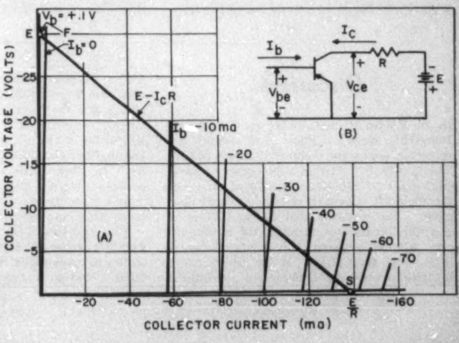


Fig. 1. (A) Collector characteristics of PNP transistor. (B) Common emitter connection showing load circuit

Windings 1 and 4 act as control windings for the switches, and the output, a square wave, appears on winding 5. To follow this operation in detail, assume that transistor *A* is conducting, which places the input voltage V_{in} across winding 2, positive at the top and negative at the bottom. Voltages with corresponding polarities will be induced in all of the other windings. This induced voltage in winding 1 makes the base of *A* negative with respect to its emitter, and *A* conducts with very little voltage drop between the emitter and collector. The voltage induced in winding 4 is positive at the top and thus holds transistor *B* cutoff, since its base is positive with respect to both the emitter and collector. Transistor *A* continues to conduct until the flux reaches the saturation value $+\phi_m$.

Upon saturation, several effects occur simultaneously; the current rises sharply to a value determined by the base drive of the transistor, the induced voltage in the transformer windings disappears, and the voltage across the transistor rises. The loss of the base drive causes transistor *A* to open the circuit to winding 2 and causes the flux to fall back slightly to the retentivity value which induces a small voltage of the opposite polarity in the windings. This places a negative voltage on the base of the transistor *B* which begins to conduct, causing the flux to decrease still further, and the cycle is rapidly reversed until transistor *B* is conducting heavily and transistor *A* is open. The input voltage is now connected across winding 3, positive at the bottom and negative at the top. The flux decreases linearly until the core saturates in the opposite direction and a similar reversal takes place to complete the cycle.

The output voltage is a square wave where the duration of each half-cycle is

---

Paper 55-73, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 25, 1954; made available for printing December 6, 1954.

G. H. ROYER is with the Westinghouse Electric Corporation, Pittsburgh, Pa.

Vol
Pt.
195

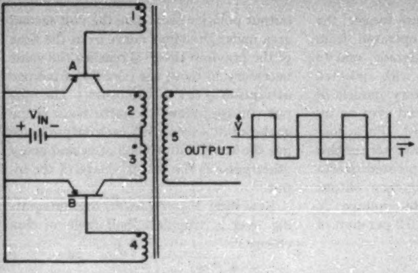
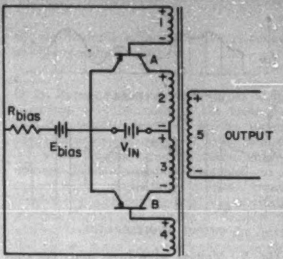


Fig. 2 (left). Basic circuit of voltage-to-frequency converter

Fig. 4 (right). Bias circuit added to increase voltage range of the basic converter



l, 74
1
55

determined by the time necessary for the flux in the core to change from $-\phi_m$ to $+\phi_m$. The frequency of these square waves can be related to the parameters of the core and the input voltage. If N_2 designates the number of turns on each one of the two main windings 2 and 3 and the voltage drop across the transistors and winding resistance is assumed to be negligible, the input voltage V_{in} is applied directly across a winding of N_2 turns

$$V_{in}=N_2\frac{d\phi}{dt}$$

Integrating this relation over one half-cycle

$$\int_0^{\frac{1}{2f}} V_{in}dt=\int_{-\phi_m}^{\phi_m} N_2 d\phi$$

Therefore

$$f=\frac{V_{in}}{4N_2\phi_m}$$

From this expression it is evident that, for a given transformer, the output frequency is proportional to the d-c input voltage. Such a characteristic is well suited for telemetering applications where the output frequency transmitted between a measuring point and a metering station is an indication of a voltage measurand at the originating point. As the transistors are used as switches or on-off devices, and their conduction resistance is only a few ohms, their characteristics do not materially affect the proportionality. The saturation flux of Hipernik V changes about 0.07 per cent per degree centigrade and, thus, the constant of proportionality will vary correspondingly with temperature.

Bias Circuit

To insure good switching operation, the minimum voltage at the transistor bases should be about 1 volt. If a 20-to-1 operating range is desired, the voltage induced would be 20 volts at the upper end of the range. The switching current in the bases would vary over a much greater than 20-to-1 range because of the nonlinear characteristic of the emitter-base diode. Consequently, excessive heating or burn-out of the transistor could occur at high currents and, if the range were extended at the low end, improper or no switching would occur at low currents.

To maintain a more satisfactory base drive over a wide range of input voltages, a current bias may be used in the base circuit, as shown in Fig. 4. To explain the operation of this circuit, E_{bias} is assumed to be much greater than the voltages induced in windings 1 or 4, so that the bias circuit effectively acts as a constant-current source of magnitude E_{bias}/R_{bias}. It is also assumed that, initially, all of this bias current is flowing through the base of transistor A. A then conducts and connects the source V_{in} to winding 2 which thereby induces voltages of the polarities shown in 1 and 4 and, e.g., with a magnitude of perhaps 1 volt. The voltage applied to the base of transistor B, V_{bB}, is equal to the sum of the voltages induced in windings 1 and 4 minus the base voltage of A, V_{bA}. A typical value for the drop across the conducting transistor A might be 1/2 volt so that the value of V_{bB}, in this example, would be $V_{bB}=1+1-0.5=1.5$, with the base positive with respect to the emitter and, hence, the base of B is cutoff and the entire bias current must be flowing through A as was originally postulated. The values of the voltage drop V_{bA} and the voltage induced in winding 1 have practically no effect on the magnitude of the bias current I_{bias} since it was assumed that E_{bias} was large. Strictly speaking, $I_{bias}=(E_{bias}+V_1-B_{bA})/R_{bias}$. Thus, in this circuit, the function of the voltages induced in the switching windings is not primarily for producing a current in the base circuit but rather for furnishing a cutoff bias for one or the other of the transistors. But, for this slight change of concept, the over-all circuit operation is identical with that already described. By using this bias circuit with a sufficiently high bias battery, the bias drive can thus be held relatively constant if desired, and this factor is no longer a limitation on the upper value of supply voltage V_{in}. Rather, the upper limit on the input voltage is set by the collector voltage that can be applied to the nonconducting transistor. Since the induced voltage in the transformer adds to the input supply voltage, the rated collector breakdown voltages must be at least twice the input voltage. In addition, the applied voltage must not cause thermal runaway of the nonconducting transistor at elevated temperatures.

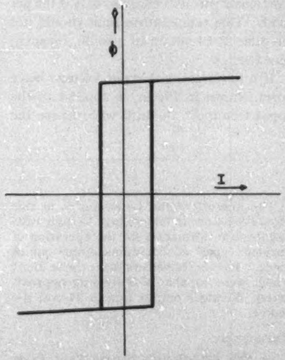


Fig. 3. Typical ϕ-I curve for square hysteresis loop core material

The low limit of V_{in} is set by the fact that, in order for transistor B to be completely cut off, the voltage V_{bB} must be at least 1/10th volt and, therefore, in this example, the sum of the voltage induced in windings 1 and 4 must be at least 1/10th volt higher than the forward drop of transistor A, V_{bA}. Thus, the greater the turns ratio N_1/N_2, the lower the volt-

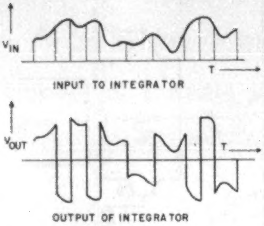


Fig. 5. Voltage wave forms using the converter as an integrator

age limitation on V_{in}. If N_1/N_2 is made too large, the base current referred to winding 2 may be greater than the maximum saturation current, and there will be nothing left over to magnetize the core so that this lower limit cannot be extended indefinitely. As the switching voltage becomes too low to saturate the transistors properly, they can no longer supply the magnetization current for the transformer. Under these conditions, the operation takes place along a minor hysteresis loop. Thus instead of acting as a saturating core, it behaves as a linear inductance in parallel with its distributed capacity. The device then changes its mode of oscillation and operates as a quasi-linear oscillator whose frequency increases as the input voltage decreases. Two precautions must therefore be observed if this device is to be used for instrumentation purposes: The input voltage must not change sign and, furthermore, it should not fall below the critical value at which the mode of oscillation changes. Input voltages of as low as 100 millivolts and up to 40 volts have been applied and, by the proper choice of cores and number of turns on windings 2 and 3, converters have operated at frequencies from one half-cycle per second up to 12 kilocycles.

Since the transistors are operating as switching devices, they will operate over a wide ambient temperature range; the circuit shown has been operated from −60 to +40 degrees centigrade, and to +75 degrees centigrade with selected transistors. Some laboratory models of this converter have operated over an input voltage range of 100 to 1. For a constant core temperature, converters of this configuration have been operated over a 20-to-1 range with a frequency output proportional to the input voltage to within a deviation of only 0.2 per cent of full-range frequency.

Applications

The circuit described in the previous section is accurate enough to warrant application of the device as a telemetering transmitter. The output square wave can be transmitted directly, or filtered, if elimination of the higher harmonics is desired. While the converter may be loaded to deliver as much as 1 watt of output power, the voltage-to-frequency linearity is sacrificed so that an output buffer stage may be necessary to maintain the 0.2-per-cent accuracy. A further advantage of a buffer stage would be that it could act as a clipper, giving a constant output magnitude over the operating range instead of one which varied with the frequency. This output can be transmitted over any desired link to an indicating station giving a representation of the voltage measured V_{in}.

It was stated that the saturation flux of Hipernik V changes about 0.07 per cent per degree centigrade and that this would vary the constant of proportionality in $f=KV_{in}$. If the core were enclosed in a constant temperature box held to a variation of ±2.5 degrees centigrade, the proportionality would change by only ±0.2 per cent. This temperature limit should not be difficult to obtain in a small, inexpensive oven.

If a unidirectional input voltage wave form, shown in Fig. 5, is applied to the input terminals, the unit will reverse the output polarity each time the volt-second area under the input curve from the time of the previous reversal reaches the value necessary to move the core flux from one saturation point to the other. The output voltage follows the same wave shape as the input, except that each output pulse has the same volt-second area and every other pulse is the mirror image of the input.

Rewriting $V_{in}=N(d\phi/dt)$ and integrating over a complete half-cycle of flux change

$$V_{in}dt=N\int_{\phi_m}^{\phi_m}d\phi=2N\phi_m$$

Thus, twice the number of turns times ϕ_m determines the volt-second area of the input which is required for each output pulse. This is then an analogue-to-digital converter in that the digital output (number of pulses) is the representation of the analogue input (integral of the applied unidirectional voltage). A count of the number of output pulses during a fixed time period would give the integral of the applied voltage over that time.

Conclusions

In this paper a simple and reliable circuit has been illustrated to convert direct to alternating current. Several of its applications—precision telemetering and integrating—have been described. Little or no effect on the operation is noted when transistors are replaced or are interchanged. It was emphasized that junction transistors used as switches make useful circuit components which will operate over widely varying ambient conditions.

References

1. Future Application of Transistors to Control Machine Tools, **R. Louis Bright.** *Proceedings*, 18th Annual Machine Tool Electrification Forum, Buffalo, New York, Apr. 1954.

2. Junction Transistors Used as Switches, **R. Louis Bright.** *AIEE Transactions*, vol. 74, pt. I, Mar. 1955, pp. 111–21.

Vol.
Pt. I
195

◆

Discussion

George C. Uchrin (Signal Corps Engineering Laboratories, Fort Monmouth, N. J.): A circuit similar to the one described in the paper was devised in January 1954 by W. O. Taylor and the writer.[1] The basic circuit and mode of operation are the same. It was found that grounded base as well as grounded emitter could be used. It was also found that almost any type of core material could be used and still obtain good wave form.

The interest of the Laboratories in this circuit is its use in low-voltage to high-voltage d-c transformation for the operation of various types of communications equipment. Power converters have been built which were capable of delivering approximately 60 watts output from a 24-volt d-c source.

Reference

1. A New Self-Excited Square-Wave Transistor Power Oscillator, **George C. Uchrin, Wilfred O. Taylor.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 43, Jan. 1955 p. 99.

A. G. Milnes (Carnegie Institute of Technology, Pittsburgh, Pa.): The switching transistor oscillator described by Dr. Royer is of considerable interest as a supply source for the operation of magnetic-amplifier equipment. The efficiency is high, the source impedance is suitably low, and the reliability is likely to be good when high-power conservatively rated transistors are available. Such converters may therefore tend to replace rotating electric machines supplying small amounts of 400-cycle-per-second power in various military applications.

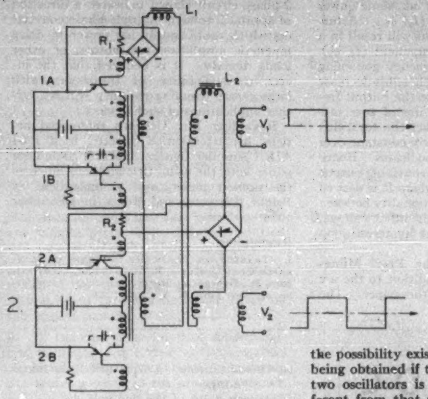


Fig. 6. Arrangement of switching-transistor oscillators

If magnetic-amplifier equipments are supplied from independent transistor oscillators it is probable that design frequencies will increase since performance for a given weight is thereby improved. This is true up to at least 2,400 cycles per second, which happens to be a frequency for which appreciable experience has been gained in magnetic-amplifier design and in the operation of associated apparatus such as motors, gyros, and torque relays.

Two switching-transistor oscillators of the same frequency can be interlinked so as to provide outputs which differ in phase by 90 degrees in a prescribed sequence. One possible arrangement is shown in Fig. 6. The transistor oscillators *1* and *2* are of the general type considered by Dr. Royer and may be operated from the same d-c source although for clarity in the diagram two separate batteries are shown. The capacitors across transistors *1B* and *2B* are not concerned with phase locking but may be desirable in the basic circuits to suppress abnormal high-frequency oscillations when the circuits are first energized. The phase-locking circuits are those containing the saturable reactors L_1 and L_2 and the rectifier bridges which apply voltages to the resistors R_1 and R_2. It will be seen that the L_1, R_1 circuit is energized by the sum of the oscillator voltages and the L_2, R_2 circuit by the difference of the voltages.

If the output voltages have a phase difference of exactly 90 degrees, the sum and difference wave forms have pulses of the same mean voltage-time area. The reactors L_1 and L_2 are designed so that the full nonsaturating flux changes occur in them for this condition and so no appreciable voltages appear across R_1 and R_2. If the phase difference between V_1 and V_2 is, however, e.g., less than 90 degrees (V_2 lagging on V_1), the pulse width of V_1+V_2 exceeds the saturation pulse width of L_1 and a residual pulse appears across R_1. This voltage switches off transistor *1A* and so causes reversal of oscillator *1* earlier than normal and therefore increases the phase difference towards the required 90-degree value. If the phase lag exceeds 90 degrees the correction is provided by the phase-locking circuit L_2, R_2 acting on oscillator *2*. Theoretically the possibility exists of no correcting action being obtained if the phase displacement of two oscillators is exactly 180 degrees different from that required, but in practice the oscillators have sufficient slip to ensure that this condition is unstable and presents no difficulty.

A rather similar circuit has been devised in which corrective action is applied to both transistors of the two oscillators and is now under examination. This uses no more components than the circuit illustrated although more windings are required.

The necessary saturation flux swings or voltage-time ratings of reactors L_1 and L_2 are solely related to the rating of the main cores and so the phase-locking action is not disturbed by change in the input voltage to the system. If the voltage-time rating of reactor L_1 is made half of that of L_2 (with the total absorption of L_1 and L_2 adjusted to match the main core), the phase lag obtained is 120 degrees. In this way with three oscillators it is possible to produce a 3-phase supply system. The prospect of obtaining 2-phase and 3-phase systems from the oscillators increases their attractiveness as supply sources.

G. H. Royer: Many different types of applications may be made of the basic d-c to a-c converter described in this paper. Mr. Uchrin's application was in the field of d-c power supply circuitry where a low-voltage direct current is converted to a square-wave alternating current transformed to a high-voltage level and rectified, yielding a good high-voltage d-c source requiring very little filtering because the rectified square wave has practically no ripple.

In such an application the only requirements are that the output voltage wave form be square to reduce the filtering required, and that the regulation of the transformer be good in order to maintain good voltage regulation. However, Prof. Milnes' discussion relates to the use of the device to supply a-c power to a magnetic amplifier or other a-c load, thus eliminating the necessity of operating at a frequency fixed by the available line power. In such an application it is desirable to maintain the output frequency constant under varying loading conditions.

The use of ordinary transformer steel as suggested by Mr. Uchrin has the obvious advantage that it is less expensive than square-loop material; however, it would not be satisfactory in Prof. Milnes' application because when such steel is used the frequency shifts considerably with a change in load. To illustrate the reason for this behavior, consider the operation of a circuit similar to that shown in Fig. 2 combining transistors and ordinary transformer steel. For a fixed input voltage V_{in} the base drive voltage to transistor *A* will be $V_{in}(N_1/N_2)$. The characteristic of transistor *A* is such that it will have a very low voltage drop as the current increases up the value I_x determined by the transistor current gain. Beyond this current the transistor voltage drop increases sharply for little further increase in current. A typical normal $\phi-i$

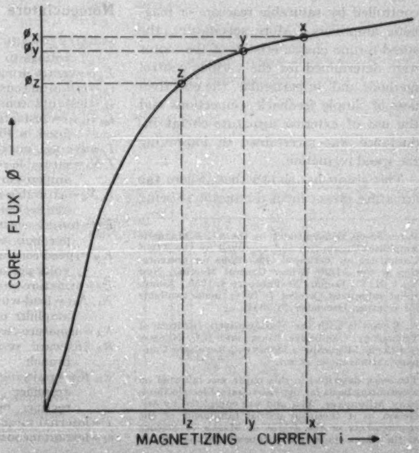


Fig. 7. Normal ϕ-i curve for transformer steel

curve for transformer steel is shown in Fig. 7. As the input voltage is applied to winding 2, the flux increases linearly until it reaches ϕ_x corresponding to the point at which the magnetizing current is equal to I_x. Here the voltage drop across the transistors increases rapidly, the voltage across the transformer falls, the induced base voltage falls, and the current in the transistor is reduced from I_x toward zero in the same manner as described in the paper. The sharp wave form is determined by the sharp current limitation and voltage increase of the transistor at the current I_x. Any output load current will be reflected back into winding 2 and will be a part of this total transistor current I_x. Thus only the difference between the reflected load current and the transistor maximum current I_x will be available to magnetize the core, and switching will occur at some lower value of current and flux (I_y, ϕ_y). A further increase in load current will result in a still lower magnetizing component (I_z, ϕ_z). It is evident from the frequency governing expression $f=V_{\text{in}}/(4N\phi_m)$ that this decrease in the flux swing will cause the output frequency to increase. The use of core materials with a flat top makes the total flux swing $2\phi_m$ much more nearly constant over a wide range of loading conditions. Hence in applications where a constant output frequency is required, or where it is desired to maintain a direct proportionality between the output frequency and the input voltage, it is necessary to use square hysteresis loop core material.

The circuitry offered by Prof. Milnes should provide a useful addition to the a-c power supply discussed in this paper. The 2-phase circuit shown indicates a direction of approach toward multiple-phase converters which could supply a-c power to drive magnetic amplifiers, a-c motors, or other loads directly. It is believed that the increased availability of transistors with large power handling capacity will see further applications of such devices.

Regarding Mr. Uchrin's referenced article, his attention is directed to a 1954 AIEE Summer General Meeting conference paper with the same title and substance as the present paper, and to an article by Bright, Pittman, and Royer in December 1954.[1]

REFERENCE

1. TRANSISTORS AS ON-OFF SWITCHES IN SATURABLE-CORE CIRCUITS, R. Louis Bright, C. F. Pittman, Jr., George H. Royer. *Electrical Manufacturing*, New York, N. Y., Dec. 1954.

Magnetic-Amplifier Control of D-C Motors

A. KUSKO
MEMBER AIEE

J. G. NELSON
NONMEMBER AIEE

ALTHOUGH considerable work has been done on the operation of adjustable-speed d-c motors with armature power supplied by thyratrons[1] and by metallic rectifiers controlled with adjustable transformers,[2,3] there is little information available on the use of magnetic amplifiers for this purpose. For this reason, an analytical and experimental study was made of the steady-state operation of a d-c motor with armature power supplied through metallic rectifiers and controlled by saturable reactors or magnetic amplifiers. More specifically, the speed-torque characteristics of the motor were determined for the various control methods and, in particular, the effectiveness of simple feedback connections and the use of external armature-circuit inductance was ascertained in improving the speed regulation.

This study has shown that, where the armature power for a d-c motor is being supplied by metallic rectifiers, adequate speed regulation and control can be obtained by the use of magnetic amplifiers with simple feedback circuits. This is particularly true where the motor speed must be adjusted remotely or by the resultant of two or more control signals. Moreover, the speed-torque characteristics of the motor and the design parameters of the auxiliary apparatus are relatively straightforward to predict.

Nomenclature

$a=E_g/V_m$ = ratio of armature generated voltage to peak line voltage
E_g = average armature generated voltage
i_A = instantaneous armature current
i_L = instantaneous line current
i_a, i_{I}, i_{II} = instantaneous currents as defined in Fig. 14
I_f = average motor field current
I_cN_c = saturable-reactor control-winding ampere-turns per core, average
I_LN_L = saturable-reactor load-winding ampere-turns per core, average
K_T = torque constant of the motor, torque per unit armature current
K_N = speed constant of the motor, generated volts per unit speed
L_A = armature-circuit inductance
N_{I}, N_{II} = load-winding turns on magnetic-amplifier cores I and II
R_A = armature-circuit resistance
R_R = forward resistance of each rectifier branch
R_{I}, R_{II} = resistance sum of one magnetic-amplifier load winding and one rectifier branch
T = internal torque of the motor
v_l = instantaneous line voltage
v_A = instantaneous armature terminal voltage
V_m = peak value of the line voltage
X_1, X_2, X_3, X_4 = branches of the bridge rectifier
α = firing angle at which core I saturates
α_1 = angle at which core I begins to absorb voltage
α_2 = angle at which rectifier X_3 begins to conduct
α_3 = angle at which the load-winding current i_{I} becomes zero
ϕ = flux reset on the magnetic-amplifier cores
$\theta=\tan^{-1}\dfrac{\omega L_A}{R_A}$ = impedance angle of the armature circuit

Analysis

The system to be considered consists of a d-c motor with constant field current, a full-wave metallic rectifier for supplying the armature power, and a saturable magnetic device for controlling the armature voltage. With respect to this system the purpose of this section is to review briefly the effect of armature-circuit inductance on the continuity of the armature current, to show what characteristics can be anticipated from control by saturable magnetic devices, and to indicate how the

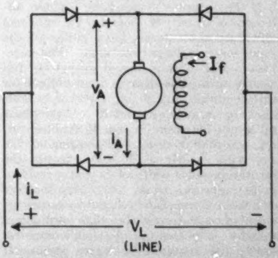


Fig. 1. D-c motor with armature power supplied by a full-wave rectifier

---

Paper 55-60, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 7, 1954; made available for printing December 3, 1954.

A. KUSKO is with the Massachusetts Institute of Technology, Cambridge, Mass., and J. G. NELSON is with the Minneapolis-Honeywell Regulator Company, Minneapolis, Minn.

The work described in this paper was initiated on a consulting basis for the Hevi Duty Electric Company, Milwaukee, Wis., and was completed by Mr. Nelson as a thesis in partial fulfillment of the requirements of the degree of Master of Science at the Massachusetts Institute of Technology.

characteristics and the design parameters can be calculated. In the analysis of these systems the following simplifying assumptions are made:

1. The armature speed and generated voltage remain constant over each cycle of armature current pulsation.

2. The metallic rectifiers can be represented electrically by a constant forward resistance and an infinite reverse resistance.

3. The hysteresis loop of the core materials used in the magnetic-amplifier and saturable-reactor cores has zero coercive force and infinite permeability below saturation and zero permeability above saturation.

4. Any voltages induced in the control windings of the magnetic amplifier or saturable reactor have negligible effect on the control current.

5. In the case of magnetic-amplifier control, inductance is added to the armature circuit of sufficient magnitude to maintain continuous current in the armature circuit.

6. In the case of magnetic-amplifier control, the armature current is zero at the instant either one of the magnetic-amplifier cores saturates. This assumes that the inductance connected in series with the armature, as discussed under item 5, is just sufficient to maintain continuous conduction.

7. In the case of magnetic-amplifier control, the armature current consists of only two well-defined periods per half-cycle of line voltage and that the circuit switches from one period to the other at $\omega t=\pi$ of the line voltage.

These last two assumptions are directly associated with the analysis of the magnetic-amplifier circuit given in the Appendix and are restated and discussed in more detail in that section.

Operation of a D-C Motor with Armature Voltage Supplied by a Rectifier

The operation of a d-c motor with armature power supplied by a full-wave rectifier, shown in Fig. 1, is basic to this discussion and will be reviewed briefly. In this system, as is well known, the armature current is pulsating. For zero armature inductance, the armature current i_A exists in pulses only at those times when the magnitude of the line voltage v_l exceeds the armature generated voltage E_g. However, when armature inductance is present, the armature current exists outside this interval; typical wave forms of armature voltage v_A and armature current i_A are shown in Fig. 2 for two extremely different values of armature inductance. The wave forms of Fig. 2(A) are typical with only the inductance of the armature itself being considered. Fig. 2(B) illustrates the effect of considerably increasing the armature-circuit inductance by adding an inductor in series with the motor armature. From these diagrams it is apparent that there are two possible modes of operation for a d-c motor supplied from a full-wave rectifier circuit depending on the relative inductance of the armature circuit. For the operation illustrated in Fig. 2(A), the armature current appears in discrete pulses, there being two such pulses during each cycle of the supply frequency. This condition is known as discontinuous conduction. The second type of operation, known as continuous conduction, is illustrated by the wave forms in Fig. 2(B). The armature-circuit inductance in this case is of sufficient magnitude to maintain current in the armature circuit at all times.

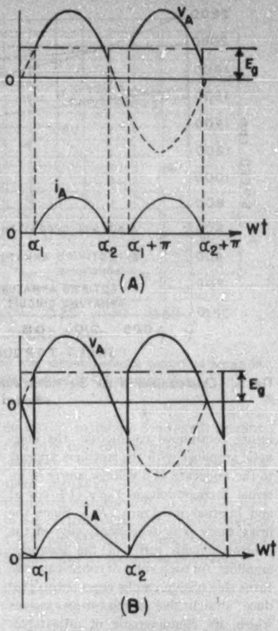


Fig. 2. Wave forms of armature current and armature terminal voltage for (A) discontinuous conduction and (B) continuous conduction

Saturable Reactor Control of Armature Current

The circuit for a saturable reactor as used to control the average current applied to the rectifier-armature circuit is shown in Fig. 3. A saturable reactor is inherently a constant-current device to the extent that, in the region of proportional-current operation, the average control-winding ampere turns I_cN_c and half-cycle average load-winding ampere turns I_LN_L in each core are equal[4]

$$I_LN_L=I_cN_c \tag{1}$$

When the motor is operating in the region of discontinuous conduction, the half-cycle wave forms of the armature and reactor load-winding currents are nearly identical so that the average values over that interval can be considered equal

$$I_A=I_L \tag{2}$$

The internal torque of the motor operating with constant field current is

$$T=K_TI_A=K_T\frac{N_c}{N_L}I_c \tag{3}$$

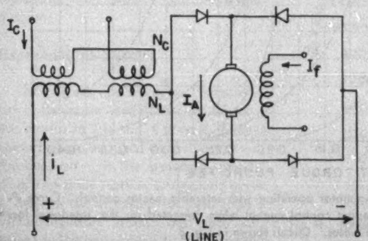


Fig. 3 (left). Circuit for saturable-reactor control of armature current

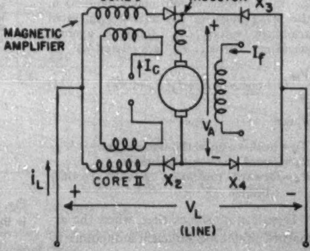


Fig. 4 (right). Circuit for magnetic-amplifier control of armature voltage

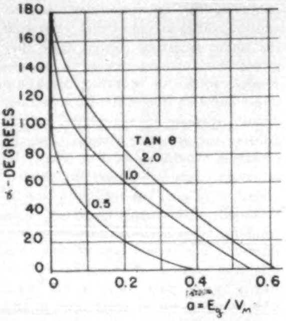


Fig. 5. Plot of equation 6 of α as a function of a for three values of tan θ

It is thus determined by the average control current and is theoretically independent of speed. When the motor is operating in the region of continuous conduction, the average armature current becomes greater than the half-cycle average value of the line current so that equation 3 is no longer valid. For adjustable-torque operation, the region of continuous conduction is of no interest.

Magnetic-Amplifier Control of Armature Voltage

A self-saturating magnetic-amplifier circuit is shown in Fig. 4 for controlling the voltage applied to the armature of a d-c motor; the circuit consists of two separate magnetic-amplifier cores I and II and four metallic rectifiers X_1, X_2, X_3, and X_4 connected in a bridge circuit. A complete analysis of this circuit is given in the Appendix.

Two important relationships stem directly from this analysis. They are as follows:

1. If the armature current is continuous, the firing angle α of the magnetic amplifier is practically independent of the speed of the motor and depends solely on the control-winding ampere turns.
2. In the region of continuous conduction, the firing angle α, the average generated armature voltage E_g, and the average armature current I_a, are related by

$$\frac{V_M}{\pi}(1+\cos\alpha)=E_g+I_A(R_A+2R_R) \tag{22}$$

where

V_M = peak value of the line voltage
R_A = armature-circuit resistance
R_R = forward resistance of each rectifier branch

Equation 22 shows that, when the armature circuit has sufficient inductance to insure continuous conduction, the magnetic amplifier with the rectifiers appears to the armature as a voltage source of internal average voltage $V_M/\pi\ (1+\cos\alpha)$ and internal resistance $2R_R$. Since the firing angle α is dependent only on the control ampere turns of the magnetic amplifier, for each value of control ampere turns this system can be expected to produce straight-line speed-torque curves which are characteristic of adjustable-voltage systems and can be predicted as easily as those for the motor alone.

The magnitude of armature inductance required to maintain continuous conduction can be determined by using the current equations derived in the Appendix. The armature current consists of two separate periods during each half-cycle of the supply frequency. These currents are given in the Appendix as

$$i_A=i_1=\frac{V_M}{(R_A+2R_R)}\left(\cos\theta\sin(\omega t-\theta)-a+[a-\cos\theta\sin(\alpha-\theta)]\epsilon^{-\frac{\omega t-\alpha}{\tan\theta}}\right)_{\omega t=\alpha}^{\omega t=\pi} \tag{24}$$

$$i_A=i_a=\left(I_0+\frac{E_g}{(R_A+2R_R)}\right)\epsilon^{-\frac{\omega t-\pi}{\tan\theta}}-\left.\frac{E_g}{(R_A+2R_R)}\right|_{\omega t=\pi}^{\omega t=\pi+} \tag{25}$$

where I_0 is the value of the armature current in equation 24 at $\omega t=\pi$. For continuous conduction it is necessary that $i_A=0$ at the instant $\omega t=\pi+\alpha$. Setting $i_A=0$ for $\omega t=\pi+\alpha$ and rearranging to a dimensionless form, equation 25 becomes

$$\frac{(R_A+2R_R)I_0}{V_M}=a\left(\epsilon^{\frac{\alpha}{\tan\theta}}-1\right) \tag{4}$$

The expression $(R_A+2R_R)I_0/V_M$ is determined by letting $\omega t=\pi$ in equation 24

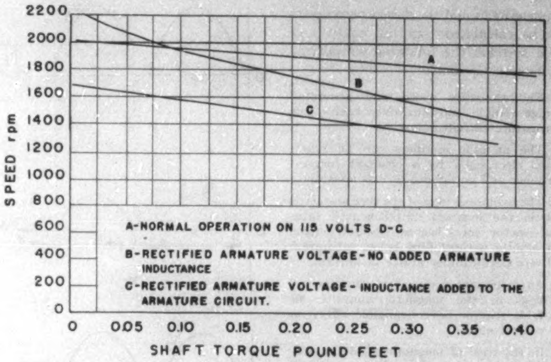


Fig. 6. Characteristics of the d-c motor without external control and with armature power supplied as shown

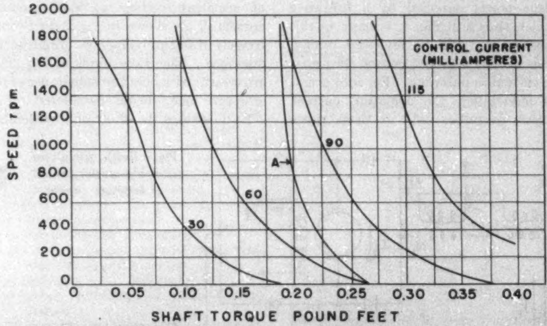


Fig. 7. Characteristics of the d-c motor operating with saturable-reactor control. Curve A is the characteristic for 60-milliampere control current when corrected for the rotational loss of the motor. Circuit shown in Fig. 3

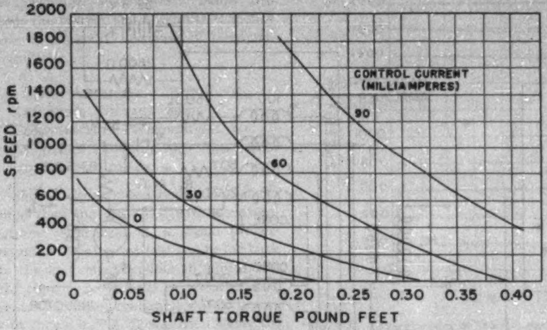


Fig. 8. Characteristics of the d-c motor operating with saturable-reactor control as shown in Fig. 3 but with an inductance of 0.20 henry added to the armature circuit

and rearranging to a dimensionless form

$$\frac{I_0(R_A+2R_R)}{V_M}=\cos\theta\sin(\pi-\theta)-a+\left[a-\cos\theta\sin(\alpha-\theta)\right]\epsilon^{-\frac{\pi-\alpha}{\tan\theta}} \qquad (5)$$

Equating equations 4 and 5 yields

$$a\left(\epsilon^{\frac{\alpha}{\tan\theta}}-1\right)=\cos\theta\sin(\pi-\theta)-a+\left[a-\cos\theta\sin(\alpha-\theta)\right]\epsilon^{-\frac{\pi-\alpha}{\tan\theta}} \qquad (6)$$

Equation 6 expresses the condition which must be met for continuous conduction in terms of a, α, and θ. A family of curves can be constructed from equation 6 to be used for determining the magnitude of armature-circuit inductance required to maintain continuous conduction. Because of the limits set on equation 4 and 5, this inductance will be that just required to maintain conduction throughout the cycle for a particular value of α, θ, and a. A graph of this form is shown in Fig. 5 where equation 6 is plotted for three different values of armature-circuit impedance angle θ($\tan\theta=0.5$, 1.0, and 2.0). The region to the right of any particular curve represents the region of discontinuous conduction while the one to the left represents the region of continuous conduction.

Experimental Results

The normal speed-torque characteristics of the motor used in this investigation are shown in Fig. 6. Curve *A* denotes the armature supplied from the 115-volt d-c line, curve *B* the armature supplied from a metallic-rectifier bridge, and curve *C* denotes the armature supplied from a metallic-rectifier bridge but with sufficient inductance added to the armature circuit to maintain continuous conduction; the field current in all three cases is constant. Note the improvement in the speed regulation shown in curve *C* as compared with curve *B*. The motor used in all of this work is a 1/8-horsepower 1,725-rpm 115-volt 1.5-ampere d-c shunt motor. The motor has no interpoles.

Saturable-Reactor Control of Armature Current

The speed-torque characteristics of the motor for the saturable-reactor control circuit shown in Fig. 3 are given in Fig. 7 with reactor control current as the parameter. The saturable reactor is a 3-legged unit, rated 220 volts, 3.0 amperes. The values of shaft torque on the abscissa are the measured brake-torque values. The internal torque is calculated as the sum of the shaft torque and the combined rotational loss torques of the motor, brake, and tachometer. This correction is made for the 60-milliampere control-current curve and is shown as curve *A* in Fig. 7.

The curves shown in Fig. 7 illustrate the adjustable-torque characteristic obtained with saturable-reactor control. At low speeds, constant torque is no longer maintained for fixed values of control current because of the onset of continuous conduction. Under conditions of continuous conduction, the average armature current exceeds the average line current; this results in an increase in torque with a decrease in speed. As an illustration of the effect of continuous conduction on the constant-torque characteristics, the curve shown in Fig. 8 was measured using the circuit shown in Fig. 3 with an inductance of 0.20 henry connected in series with the armature. This inductance was sufficient to maintain continuous conduction over the entire lower range of motor speeds producing the pronounced departure of these curves from the ideal constant-torque characteristic. For this reason, armature-circuit inductance must be kept to a minimum if the torque is to be a function only of the control current and not of the speed in the circuit shown in Fig. 3.

Magnetic-Amplifier Control of Armature Voltage

Experimentally determined speed-torque characteristics for the magnetic-amplifier bridge circuit shown in Fig. 4 are given in Fig. 9; shaft torque is the measured brake torque. The magnetic amplifier used in this work is a commercially available 60-cycle-per-second unit, rated 1.1 amperes, 180 watts.

The relationship between average applied voltage, motor internal torque, and speed is derived in the Appendix and is

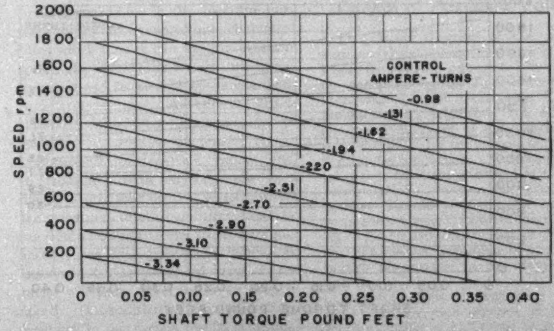


Fig. 9. Characteristics of the d-c motor operating with magnetic-amplifier control, as shown in the circuit of Fig. 4

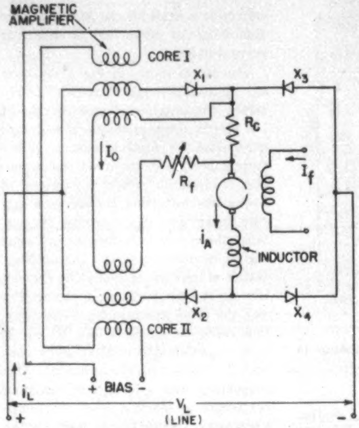


Fig. 10. Circuit for magnetic-amplifier control of the d-c motor with compensation proportional to armature current

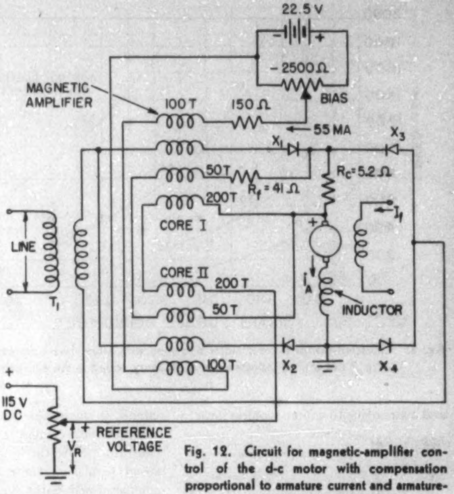


Fig. 12. Circuit for magnetic-amplifier control of the d-c motor with compensation proportional to armature current and armature-voltage feedback

$$n=\frac{1}{K_n}\left(\frac{V_M}{\pi}(1+\cos\alpha)-\frac{T}{K_T}(R_A+2R_R)\right) \tag{23}$$

To verify the validity of equation 23, the experimental curves must be corrected for rotational loss to obtain internal torque. Consider, e.g., the curve for 0.98 control ampere-turns. When this curve is corrected for rotational loss, the drop in speed at full load is 0.450 of the no-load speed. The per-unit resistance of the armature circuit including the resistance of the external inductor is 0.20 and the per-unit voltage drop of one rectifier branch is approximately 0.130, based on the source voltage at no-load speed and the full-load armature current. The calculated per-unit speed regulation is then 0.20+2×0.13=0.46, which agrees closely with the given experimental value.

The speed regulation obtained with this basic magnetic-amplifier circuit can be improved by decreasing the rectifier forward resistance to as low a value as possible, or by employing some type of feedback or compensation to increase the average voltage applied to the armature in proportion to the average armature current or torque as the load is increased. With the metallic rectifiers now available commercially, the forward resistance cannot be made appreciably less than that encountered in this investigation, so that feedback or compensation methods are required to obtain an improvement in speed regulation.

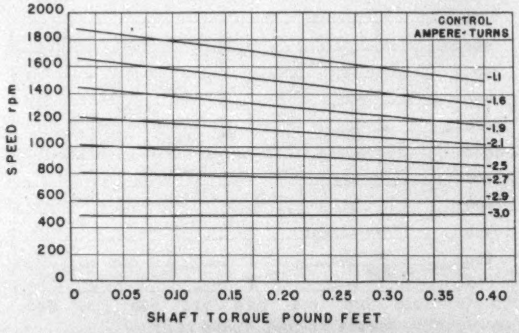


Fig. 11. Characteristics of the d-c motor operating with the magnetic-amplifier control circuit shown in Fig. 10. Compensation proportional to armature current

Magnetic-Amplifier Control of Armature Voltage with Compensation Proportional to Armature Current

Since the speed-torque characteristics shown in Fig. 9 are straight lines, the simplest method of improving the speed regulation of the circuit shown in Fig. 4 is by the use of compensation proportional to armature current, as shown in Fig. 10. In this scheme, a voltage proportional to armature current is applied to a second control winding on the magnetic-amplifier cores. The relative polarity of the resulting current I_0 is such that the magnetic-amplifier firing angle is decreased with increasing load on the motor. The speed level (no-load speed) is adjusted with the bias-control current, while the speed drop (compensation) is controlled by the compensation-circuit resistance R_f.

The speed-torque characteristics shown in Fig. 11 are an example of the type of speed regulation that can be obtained with this circuit. To obtain the data shown in Fig. 11, the compensation-circuit resistance was adjusted for flat regulation at 600 rpm. For speeds above 600 rpm, the speed regulation becomes progressively greater (more drop) although the regulation for a no-load speed of 1,900 rpm is only 0.23 per unit, which is a considerable improvement over that obtained with no

Vol. 7
Pt. 1
1955

compensation whatsoever. Below 600 rpm, the system has a definite tendency toward instability because of overcompensation.

Magnetic-Amplifier Control of Armature Voltage with Compensation Proportional to Armature Current and Armature-Voltage Feedback

Should closer speed regulation be desired than that obtained with armature-current compensation, a signal proportional to armature voltage can be fed back to the magnetic amplifier in addition to the signal proportional to armature current. In the experimental work, the circuit shown in Fig. 12 was used where the three control windings on the magnetic-amplifier core serve as a bias winding, compensation winding, and feedback winding. The feedback-winding signal is proportional to the difference between the armature terminal voltage and a reference voltage, so that this is a true feedback control system. The motor speed is adjusted by means of the reference voltage V_R. An isolating transformer T_1 was used in the experimental work so that one side of the armature circuit could be grounded to the particular d-c source used for reference.

A set of experimentally determined speed-torque curves for this control scheme is illustrated in Fig. 13 where speed and torque are shown for parametric values of reference voltage. The circuit parameters and number of control-winding turns used in obtaining these characteristics are given in the diagram shown in Fig. 12. These curves illustrate the characteristics that can be obtained with magnetic-amplifier control of armature voltage when adequate compensation and feedback are used. The system is stable at all speed levels and is critically damped or slightly overdamped in response to a change in either reference voltage (speed) or load. No definite data were taken regarding response time.

The experimental characteristics shown in Fig. 13 were measured with continuous conduction in the armature circuit. For this condition no stabilization or filtering was necessary in any of the control windings to maintain a stable system. With the external inductor removed so that conduction was discontinuous, the system became unstable. Inductors of approximately 10 henrys were added to each of the control windings to alleviate this condition and, although this reduced the oscillations in speed to a certain extent, the system still had definite tendencies to oscillate at any set speed level. From this experience it is apparent that continuous conduction is essential for stable operation of this circuit unless further stabilization techniques are to be employed.

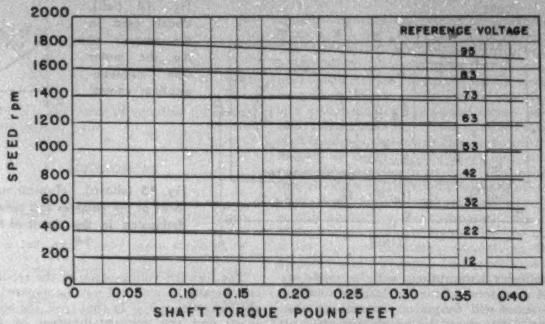


Fig. 13. Characteristics of the d-c motor operating with the magnetic-amplifier control circuit shown in Fig. 12. Compensation proportional to armature current and armature-voltage feedback

Conclusions

The control of the armature voltage of an adjustable-speed d-c motor by means of saturable magnetic devices produces well-defined torque-speed characteristics. Saturable-reactor control results in nearly constant-torque characteristics at each level of control current; magnetic-amplifier control produces approximately constant-speed characteristics at each level of control current.

Continuity of conduction of the armature current influences the shape of the torque-speed characteristics of the controlled motor. With saturable-reactor control, discontinuous conduction in the armature current results in speed-torque characteristics adhering closely to constant torque levels. On the other hand, under magnetic-amplifier control, continuous conduction in the armature current results in speed-torque characteristics which may not be exactly constant speed, but at least are straightline in nature and can be altered by simple feedback arrangements. An external armature-circuit inductor is necessary to achieve continuous conduction.

Compensation feedback proportional to armature current may be applied to a magnetic-amplifier control system, where the armature current has continuous conduction, to obtain at one speed a torque-speed curve with zero speed regulation; the remaining torque-speed curves in the range will show either a rise or drop in speed. The addition of armature-voltage feedback to this system results in torque-speed curves with close-speed regulation at all speed levels in the range.

The type of magnetic-amplifier control system described is suited to applications requiring speed regulation of the order of 2 per cent at one speed and a speed range of the order of 10-or-20 to 1. For this purpose, the system is certainly simpler than a thyratron system and is about as complex as an adjustable-transformer metallic-rectifier system where the field winding of the motor is supplied from a separate rectifier. The magnetic amplifier has the further advantage that it can be controlled remotely, at low power levels, and from multiple signals.

Further work is needed on methods for predicting the dynamic performance of such systems. Experience in the experimental work on the units described indicated that the transient response of the motor speed to changes in control current was fast and adequately damped.

Appendix I. Analysis of the Magnetic-Amplifier Circuit for Control of Armature Voltage

The circuit diagram of the magnetic amplifier and d-c motor is shown in Fig. 14 where the motor is represented by an internal generated voltage E_g in series with an inductance L_A and a resistance R_A. The inductance L_A represents the armature-winding inductance plus an external inductance added for the purpose of obtaining continuous conduction. The resistance R_A represents the total resistance of the armature winding, brushes, and external inductor. The forward rectifier resistance is assumed constant and is represented by R_R. The resistance of one load winding and the forward resistance of one rectifier branch are lumped in the resistances R_I and R_{II}.

In this analysis, the equations will first be written for each separate interval occurring during a half-cycle of operation. When this has been accomplished, necessary

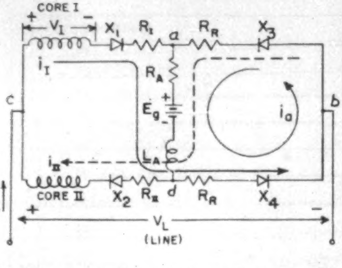


Fig. 14 (left). Circuit used in the analysis of the d-c motor with magnetic-amplifier control

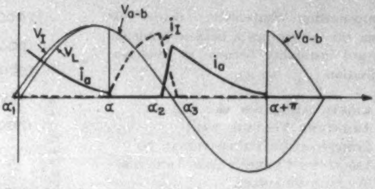


Fig. 15 (above). Typical wave forms of the voltages and currents designated in the circuit of Fig. 14

simplifying assumptions will be made so that an expression can be obtained relating the speed and torque of the motor. The approximate wave forms of currents and voltages are shown in Fig. 15, which also serves to define the various intervals of operation. The assumptions relating to this analysis are given previously in the section entitled "Analysis."

The Angle $\omega t = \alpha_1$

The angle α_1 is defined as the angle at which core I begins to absorb voltage. During the negative half-cycle of the supply voltage $v_l = V_m \sin \omega t$, the flux in core I has been reduced below saturation to a value depending on the control ampere-turns.

The currents i_I, i_{II}, and i_a are defined in Fig. 14; i_I is the current in the magnetic-amplifier load winding placed on the core I, i_{II} is the current in the load winding on core II, and i_a is a current in the loop *a-d-b*. Following the time $t = \alpha_1/\omega$, core I is absorbing voltage, the current $i_I = 0$, so that the voltage applied to core I is

$$v_I = V_M \sin \omega t + i_a R_R \tag{7}$$

where $i_a R_R$ is the voltage drop across the rectifier X_3. The flux in core I will begin to increase toward saturation when $v_I > 0$ or when

$$V_M \sin \omega t \geq -i_a R_A \tag{8}$$

The Interval $\omega t = \alpha_1$ to $\omega t = \alpha$

The angle α is defined as that angle at which magnetic amplifier core I saturates and is referred to as the firing angle. The core saturates when

$$\varphi = \frac{1}{N_I}\int_{\alpha_1}^{\alpha} v_I dt = \frac{1}{N_I}\int_{\alpha_1}^{\alpha} (V_M \sin \omega t + i_a R_R) dt \tag{9}$$

where φ is the flux reset during the preceding half-cycle and N_I are the load-winding turns on core I. During this interval, from Fig. 14

$$i_a R_R = -\left(E_g + L_A \frac{di_a}{dt} + i_a(R_A + R_R)\right) \tag{10}$$

so that the core will saturate when

$$\varphi = \frac{1}{N_I}\left\{\int_{\alpha_1}^{\alpha}\left[V_M \sin \omega t - \left(E_g + L_A \frac{di_a}{dt} + i_a(R_A + R_R)\right)\right]dt\right\} \tag{11}$$

The forward voltage drop of the rectifier can usually be assumed small compared to the line voltage v_l. In that case, the angle $\alpha_1 = 0$ and the voltage absorbed by the core in the interval α_1 to α is the line voltage v_l. The relationship between the reset flux φ and the firing angle α is then independent of voltage E_g and the speed. This only occurs because the armature current is forced to be continuous throughout the interval and rectifier X_3 remains conducting.

The Interval $\omega t = \alpha$ to $\omega t = \alpha_2$

The angle $\omega t = \alpha_2$ is defined as the angle at which the rectifier X_3 begins to conduct, following the pulse of current through load winding I; this provides a circulating path around the armature circuit for the current i_a. At $\omega t = \alpha$, core I saturates and the voltage $V_M \sin \alpha$ is applied instantaneously between points *a* and *b*; see Fig. 14. If, as will usually be the case at that instant

$$V_M \sin \alpha \geq i_a R_R \tag{12}$$

the rectifier X_3 will block and the current will commutate, so that

$$i_I = i_a \Big|_{\omega t = \alpha} \tag{13}$$

Rectifier X_3 will continue to block as long as $v_{ab} > 0$. The current in load winding I and the armature during the interval α to α_2 is then the solution to the equation

$$V_M \sin \omega t = E_g + L_A \frac{di_I}{dt} + i_I(R_I + R_A + R_R) \tag{14}$$

where the boundary condition for i_I at $\omega t = \alpha$ is given by equation 13. At the instant that

$$v_{ab} = V_M \sin \omega t - i_I R_I = E_g + L_A \frac{di_I}{dt} + i_I(R_A + R_R) = 0 \tag{15}$$

rectifier X_3 will begin to conduct allowing a current i_a to circulate around the loop *a-d-b-a*, as shown in Fig. 14.

The Interval $\omega t = \alpha_2$ to $\omega t = \alpha_3$

The angle $\omega t = \alpha_3$ is defined as the angle at which the load-winding current i_I becomes zero. After $\omega t = \alpha_2$, the equations for the currents i_I and i_a are

$$V_M \sin \omega t = E_g + L_A \frac{d}{dt}(i_I + i_a) + i_a R_A + i_I(R_A + R_I + R_R) \tag{16}$$

$$V_M \sin \omega t = i_I R_I - i_a R_R \tag{17}$$

with the boundary conditions at $\omega t = \alpha_2$ being

$$i_a = 0$$

$$i_I = I_K$$

where $i_I = I_K$ is the value of the solution to equation 14 at $\omega t = \alpha_2$.

The Interval $\omega t = \alpha_3$ to $\omega t = \alpha + \pi$

The angle $\alpha + \pi$ is the angle at which core II saturates in the negative half-cycle of line voltage. During this interval, core I is resetting and the load-winding current $i_I = 0$; the current i_a is the solution to the equation

$$0 = E_g + L_A \frac{di_a}{dt} + i_a(R_A + 2R_R) \tag{18}$$

with the boundary condition

$$i_a = \frac{-V_M \sin \omega t}{R_R}\Bigg|_{\omega t = \alpha_3}$$

This current continues until $\omega t = \alpha + \pi$ when core II fires and the foregoing sequence of currents repeats itself.

The equations pertaining to each distinct interval occurring during a cycle of operation have been established in an implicit form. The current in each of the foregoing intervals begins with a boundary value established by the preceding interval. If these boundary values were to be applied rigorously in the equation solutions which follow, the resulting expression for average load current would be too cumbersome to be of practical value. The solutions which follow, therefore, are based on two simplifying assumptions made for the purpose of obtaining a relatively practical expression for the speed and torque of the d-c motor. These assumptions are listed previously as items 6 and 7 in the section entitled "Analysis" and are discussed further at this point. These are:

1. The external inductance in the armature circuit is just sufficient to maintain continuous conduction; the load-winding current i_I and the circulating armature current i_a are both zero at the instant $\omega t = \alpha$.
2. The load winding current i_I decreases instantaneously to zero at the instant $\omega t = \pi$ and, at this instant, the armature circulating current i_a begins with the value of i_I just before $\omega t = \pi$ as the initial boundary condition. If the voltage drops in the magnetic-amplifier load winding and in the rectifiers X_1 and X_3 are assumed small; this actually is the case since then rectifier X_1 blocks at the same instant that rectifier X_3 begins to conduct.

With the foregoing two assumptions the armature current now consists of only two intervals: From $\omega t=\alpha$ to $\omega t=\pi$

$$V_M \sin \omega t=E_g+L_A\frac{di_I}{dt}+i_I(R_I+R_A+R_R) \tag{14}$$

where

$i_I=0$ at $\omega t=\alpha$

This current decreases instantaneously to zero at $\omega t=\pi$. From $\omega t=\pi$ to $\omega t=\alpha+\pi$

$$0=E_g+L_A\frac{di_a}{dt}+i_a(R_A+2R_R) \tag{18}$$

where

$i_a=i_I$ at $\omega t=\pi$

$i_a=0$ at $\omega t=\alpha+\pi$

Assuming that $R_A+R_R+R_I \simeq R_A+2R_R$ and taking the average values of the terms in equations 14 and 18 over one half-cycle

$$\frac{1}{\pi}\int_{\omega t=\alpha}^{\omega t=\pi} V_M \sin \omega t d(\omega t)=\frac{1}{\pi}\int_{\omega t=\alpha}^{\omega t=\pi} [E_g+i_I(R_A+2R_R)]d(\omega t)+\frac{1}{\pi}\int_{\omega t=\alpha}^{\omega t=\pi} L_A\frac{di_I}{dt}d(\omega t) \tag{19}$$

$$0=\frac{1}{\pi}\int_{\omega t=\pi}^{\omega t=\alpha+\pi} [E_g+i_a(R_A+2R_R)]d(\omega t)+\frac{1}{\pi}\int_{\omega t=\pi}^{\omega t=\alpha+\pi} L_A\frac{di_a}{dt}d(\omega t) \tag{20}$$

Recognize that the flux linkage excursion in the armature-circuit inductance is cyclic such that

$$\int_{\omega t=\alpha}^{\omega t=\pi} L_A\frac{di_I}{dt}d(\omega t)=-\int_{\omega t=\pi}^{\omega t=\alpha+\pi} L_A\frac{di_a}{dt}d(\omega t) \tag{21}$$

The addition of equations 19 and 20 results in a simple expression for the average armature current I_A

$$\frac{V_M}{\pi}(1+\cos \alpha)=E_g+I_A(R_A+2R_R) \tag{22}$$

The relationships for speed and torque, $n=E_g/K_n$ and $T=K_TI_A$, can be substituted in equation 22 resulting in

$$n=\frac{1}{K_n}\left(\frac{V_M}{\pi}(1+\cos \alpha)-\frac{T}{K_T}(R_A+2R_R)\right) \tag{23}$$

Equation 23 is the final expression for the speed of the motor in terms of torque and firing angle. The firing angle α is a function of control current only.

The solutions to equations 14 and 18 for the instantaneous currents are given by Vedder and Puchlowski as[1]

$$i_I=\frac{V_M}{(R_A+2R_R)}\left[\cos \theta \sin (\omega t-\theta)-a+[a-\cos \theta \sin (\alpha-\theta)]\epsilon^{-\frac{\omega t-\alpha}{\tan \theta}}\right]_{\omega t=\alpha}^{\omega t=\pi} \tag{24}$$

$$i_a=\left(I_0+\frac{E_g}{R_A+2R_R}\right)\epsilon^{-\frac{\omega t-\pi}{\tan \theta}}-\frac{E_g}{R_A+2R_R}\Bigg|_{\omega t=\pi}^{\omega t=\alpha+\pi} \tag{25}$$

where I_0 is the value of equation 24 at the instant $\omega t=\pi$.

References

1. THEORY OF RECTIFIER—D-C MOTOR DRIVE, E. H. Vedder, K. P. Puchlowski. *AIEE Transactions*, vol. 62, 1943, pp. 863–70.

2. AN IMPROVED VARIAC SPEED CONTROL, W. N. Tuttle. *The General Radio Experimenter*, General Radio Company, Cambridge, Mass., vol. XXV, no. 12, May 1951.

3. A THREE-QUARTER HORSEPOWER VARIAC MOTOR SPEED CONTROL, W. N. Tuttle. *Ibid.*, vol. XXVI, no. 12, May 1952.

4. THE TRANSDUCTER AMPLIFIER (book), U. Krabbe. Lindhska Boktryckeriet, Orebro, Sweden, 1947.

Discussion

L. A. Finzi (Carnegie Institute of Technology, Pittsburgh, Pa.): One wonders why the authors in their analytical treatment have chosen to introduce a rather artificial assumption; they assume that the external choke used in the armature circuit is just sufficient to cause the armature current to decay to zero immediately before a new firing occurs. The reader observing Fig. 5 recognizes that for any given choke there is only one angle of firing at any shaft speed for which this condition is met and he may worry about the broader use of equation 23, thus derived, to describe the general behavior of the motor.

Equation 23 is valid mathematically over the whole range of motor operation in which continuous flow of armature current prevails. This can be shown by the comparatively simple integration of the linear first-order differential equations written under the sole use of assumption 2 of the Appendix; the dubious assumption 1 is totally unnecessary. As a matter of fact, equation 23 can be written also directly from half-cyclic volt-time area considerations, similar to those that have become most usefully popular in the study of magnetic amplifiers. The half-cyclic volt-time area made available by the amplifier upon firing equals the sum of the volt-time areas of resistive drops and of armature voltage E_g (which at no time is blocked by the rectifiers, if the flow of current is continuous); the half-cyclic volt-time area absorbed by the inductances of the armature circuit obviously is zero.

The particular design described in the paper is characterized by the unusual employment of a large choke in series with the armature; hence the range of torques and speeds over which continuous conduction occurs may be broader than in the operation of practical variable-speed drives of this same type commonly marketed by various manufacturers. But in any case, modes of discontinuous conduction are not unlikely over the range. Here equation 23 no longer applies and other expressions are needed to relate α, E_g, and average armature current. The derivation of such expressions becomes less simple, and one may agree that the process of matching of boundary values would be tedious. However, workable equations and complementary graphical means have been made available.[1] This reference deals with irksome questions arising in the general loading of magnetic amplifiers and gives particular attention to the case of d-c armature circuits in all modes of current flow.

The analysis mentioned was performed to acquire knowledge about armature currents in the various modes of conduction and also about the corresponding currents in the amplifier load windings. In the problem at hand the variables of ultimate interest are the load torque, the shaft speed and the quantity chosen to control the magnetic amplifier; the angle of firing is only a convenient intermediate parameter to be eliminated in the final relationship. The need for knowing amplifier power-winding currents is most evident in the case of armature control by saturable reactors. Equation 3 is based on a presumed identity of currents in armature and amplifier load windings, as the authors recognize that no quantitative agreement should be expected with the experimental results of Fig. 7, especially in the critical region of low speeds. With a more subtle mechanism, dictated by the dynamic characteristics of the cores, the wave form of load-winding current affects the amplifier signal requirements for wanted angle of firing also in the case of many self-saturating amplifiers.[2] Whether or not these recognizable effects should be ignored depends largely on the scope of the amplifier system designer and of the problems he faces.

REFERENCES

1. THE OPERATION OF MAGNETIC AMPLIFIERS WITH VARIOUS TYPES OF LOAD. PART I—LOAD CURRENTS FOR GIVEN ANGLE OF FIRING, L. A. Finzi, R. R. Jackson. *AIEE Transactions*, vol. 73, pt. 1, July 1954, pp. 270–78.

2. PART II—CONTROLLING THE ANGLE OF FIRING. THE TRANSFER CHARACTERISTICS OF AMPLIFIERS WITH LOW-CONTROL IMPEDANCE. *Ibid.*, pp. 279–88.

A. Kusko and **J. G. Nelson**: Professor Finzi correctly points out that the equations for operation of this system were derived on the basis of armature inductance just sufficient to maintain continuous conduction, yet the results were applied over the complete speed range. As he states, the results can be obtained without this assumption. Actually, equations 19, 20, and 21 basically represent the volt-time-area approach to the problem.

Additional armature-circuit inductance was not included merely to obtain continuous conduction and a simpler analysis, but was inserted to obtain the straight-line speed-torque curves that resulted from continuous conduction. These curves then made possible the effective compensating means for improving the speed regulation within the economic considerations of the application for which the drive was studied. We acknowledge the generality of Professor Finzi's work on this type of problem and appreciate the usefulness of his references in handling problems of discontinuous conduction in such circuits.

Equations for the Inductance and Short-Circuit Forces of 3-Phase Busses Comprised of 120-Degree Angles

J. W. MAIER
ASSOCIATE MEMBER AIEE

T. J. HIGGINS
MEMBER AIEE

THREE-PHASE busses comprised of 120-degree-angle conductors, as shown in Fig. 1, possess certain very desirable electrical-mechanical characteristics. Thus, they can be closely nested to the end that the reactance per unit length of bus, and thus the voltage drop per unit length of bus, are much smaller than on other shaped busses of similar current-carrying capacity. Again, 120-degree-angle conductors have considerable mechanical rigidity with which to withstand short-circuit forces. Finally, their shape enables an open, compact construction which simplifies ventilation and heat-dissipation problems, minimizes the space required for installation, and provides easy tap-off to individual loads.

These electrical, mechanical, and physical characteristics render angle conductors especially suited for long heavy-current 3-phase distribution busses where it is desirable to reduce the reactance per unit length to a minimum and to minimize the number of insulating supports so as to reduce the number of possible flashover points. In consequence, 3-phase busses comprised of 120-degree-angle conductors have found increasing use[1-6] as a preferred type of construction for distribution busses in high-capacity generating stations and for heavy-current industrial loads.

To predetermine the operating characteristics of these busses requires knowledge of the a-c resistance, the a-c inductance, the copper temperature under load, and the mechanical forces exerted during short circuit. But, though equations are available for calculating the a-c resistance[7] and copper temperature,[8,9] convenient equations, in terms of the bus dimensions, are not available for calculating the inductance and short-circuit forces. It is obvious, however, that knowledge of these equations is most desirable to the bus designer.

Such equations are derived in this paper, it being assumed that the conductors are nonmagnetic, of such length that end effects are negligible, of uniform thickness, right-cornered, and that they carry currents distributed uniformly over their cross sections. Of these five conditions, the first three are satisfied in practice. The fourth, however, though essential to tractable analysis, is not always true in practice. Angles drawn, rolled, bent, extruded, or otherwise formed in one piece have rounded edges; but the error introduced by neglecting the rounded edges is small and, where necessary, the increase in inductance caused by these rounded edges can be approximated by use of semiempirical equations derived by Dwight and Wang.[10] Finally, though skin and proximity effects render the current density over the cross sections of the bus conductors slightly nonuniform, the analysis in the section entitled "Corrections for Rounded Edges and for Skin and Proximity Effects" of reference 11, the discussion following reference 10, and the experimental results published in references 12 and 13 all indicate that, at power frequencies, the a-c inductance is essentially identical with the d-c inductance.

If, then, the stated five postulates are granted, derivation of the desired equations, though lengthy, is not particularly difficult. The conductor cross sections can, per well-established theory discussed in the following, be considered as combinations of line segments; see Fig. 2. By virtue of the fundamental theorem of geometric mean distance theory, the inductance can be expressed in terms of the geometric mean distances between these line segments. These geometric mean distances can be calculated from known equations,[14-16] as shown under the section entitled "General Relationships." Accordingly, determination of the inductance as a function of the parameters defining the spacing and cross sections of the bus reduces, in essence, to specific formulation of these geometric mean distances and substitution of them in the general equation for the inductance, as explained in the section entitled "Derivation of the General Equation for the Inductance." To check the correctness of the resulting equation, it is noted that, if c in Fig. 2 is set equal to zero, the bus reduces to the simple configuration of Fig. 3 whose inductance is a very simple expression, easily written down. Corroboratively, this simple equation for the bus of Fig. 3 is yielded in the limit as $c \to 0$ by the general equation for the bus of Fig. 2; see section entitled "Corrobation of Equation 15."

The mechanical forces acting on busbar supports during short circuits are determined both by the magnitudes of the mutual electromagnetic forces exerted among the bus conductors and by the elastic properties of the supporting structure: its motional resistance, natural frequencies, and physical nature and arrangement. Comprehensive mathematical studies, substantiated by experimental data, of the individual and collective efforts of these factors and general methods for calculating the magnitudes involved have been advanced by various authors.[17-22] In shorter articles based on the work of Schurig,[17] certain charts, nomograms, and short cuts that facilitate the numerical labor incident to actual calculation have been presented.[23-27] In consequence of the work of these men, it may be taken that, if the designer has knowledge of the physical make-up of the bus structure and of the mutual electromagnetic forces exerted among the bus conductors, he can calculate readily the forces on the bus supports. Obviously, knowledge of the physical make-up of the bus structure is immediately available to the designer. Accordingly, if there are also available to him equations for calculating the mutual electromagnetic forces, he has all that is essential to calculate the bus support stresses.

These desired equations for the electromagnetic forces can be obtained by simple differentiation of the equations for the inductance.[28] The equations so obtained, however, require rather lengthy calculation. An alternative, simple means of approximating the desire values is advanced; see the section

---

Paper 55-126, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 19, 1953; made available for printing November 29, 1954.

J. W. Maier is with the Bell Aircraft Corporation, Buffalo, N. Y., and T. J. Higgins is with the University of Wisconsin, Madison, Wis.

This paper is based upon a thesis supervised by Prof. Higgins and submitted by Mr. Maier to the faculty of the University of Wisconsin in partial fulfillment of the requirements for the degree of Master of Science in electrical engineering.

Much of the content of this paper stems in part from a project originally suggested to Prof. Higgins by the Chase Brass and Copper Company, and is made available herein with their permission.

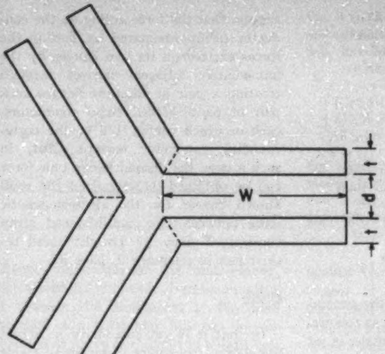


Fig. 1 (left). Cross section of 120-degree-angle bus

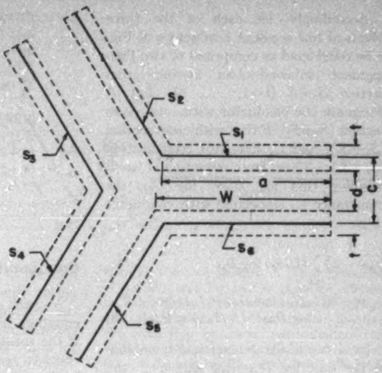


Fig. 2 (right). Cross section of 120-degree-angle line-segment bus

entitled "Equations for Electromagnetic Forces."

Finally, application of the equations for the inductance and electromagnetic forces is illustrated by numerical solution of a typical problem in bus design in the section entitled "An Illustrative Example."

General Relationships

INDUCTANCE

From well-known theory[14,15] it has been derived that the inductance associated with the kth conductor of a bus comprised of r long parallel nonmagnetic conductors, each having a uniform but otherwise arbitrary cross section (reference 15, equation 27), is

$$L_k = -2\sum_{i=a}^{r}(I_i/I_k)\log D_{ik} \text{ abhenries per centimeter of conductor length} \quad (1)$$

where D_{ik} is the geometric mean distance between the kth and ith conductors and I_i is the current (expressed as a complex number) associated with the ith conductor. Equation 1 can be written

$$L_k = \sum_{i=a}^{r} L_{ik} \quad (2)$$

where L_{ik} is the contribution of the ith conductor to the inductance associated with the kth conductor and is given by

$$L_{ik} = -2(I_i/I_k)\log D_{ik} \quad (3)$$

The currents satisfy the Kirchhoff current law

$$I_a+I_b+I_c+\ldots+I_r=0 \quad (4)$$

SHORT-CIRCUIT FORCES

A convenient method[28] for finding the electromagnetic force F_j on the jth conductor of a bus comprised of r conductors is to use the relations

$$F_{jx}=\partial W/\partial x_j \text{ and } F_{jy}=\partial W/\partial y_j \quad (5)$$

where x_j and y_j define the position of the jth conductor in a system of Cartesian co-ordinates x and y in the plane of the cross section of the bus. F_{jx} and F_{jy} are the x and y components respectively of the total force F_j on the jth conductor. W is the energy in the magnetic field of the bus and is given by

$$W=\sum_{k=a}^{r} L_k I_k^2/2 \quad (6)$$

where L_k is given by equation 1. Substituting from equation 6 in equation 5 yields

$$F_{jx}=\sum_{k=a}^{r}(I_k^2/2)\partial L_k/\partial x$$

and

$$F_{jy}=\sum_{k=a}^{r}(I_k^2/2)\partial L_k/\partial y \quad (7)$$

Derivation of the General Equation for the Inductance

The 3-phase bus of Fig. 1 is comprised of three identical conductors, to be distinguished by A, B, and C. In virtue of the geometrical symmetry it follows that the inductances associated with the three conductors are identical, thus $L_A=L_B=L_C=L$. Now attempt to calculate L by appropriate subdivision of each of the three subconductors into two identical subconductors of trapezoidal cross section, as indicated in Fig. 1, and subsequent use of equation 1 would, because of the complicated expressions for the self- and mutual geometric mean distances of these trapezoidal areas, result in an equation more or less intractable for actual numerical computation. An equation tractable to computation can, however, be obtained as follows.

Let the conductors of a bus be comprised of structural shapes such as straps, channels, tubes, angles, or the like and let the ratios of the external major cross-sectional dimensions to the thickness t approximate 10 or greater. Then, per well-established theory[10,12,30] confirmed by experiment, the inductances and electromagnetic forces associated with the actual conductors are closely approximated by those which stem from replacing the cross section of each conductor by line segments having dimensions equal to the mean dimensions of the actual conductor; see Fig. 2. Now the angles used in practice have a total flange width to thickness ratio $2W/t$ of 20 to 50; whence, the bus of Fig. 1 can be well approximated by the line-segment bus of Fig. 2.

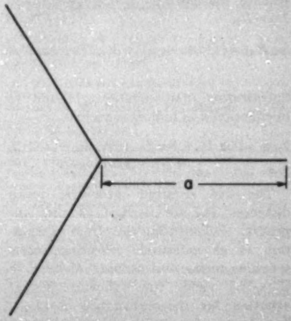


Fig. 3. Cross section of bus with zero spacing

Accordingly, let each of the three identical line-segment conductors of Fig. 2 be considered as comprised of two line-segment subconductors having cross section $S_i=S$ $(i=1, \ldots, 6)$. Let A designate the conductor with total cross section (S_1+S_2), B that with cross section (S_3+S_4), and C that with cross section (S_5+S_6). Then from equation 1, and recalling that $\log D_{ij}= \log D_{ji}$, the inductance L_A associated with conductor A is

$$L_A=-2\sum_{i=A}^{C} (I_i/I_A) \log D_{Ai}$$
$$=-2[(I_A/I_A) \log D_{AA}+(I_B/I_A)\times \log D_{AB}+(I_C/I_A) \log D_{AC}] \quad (8)$$

In virtue of the geometrical symmetry of the bus, $\log D_{AC}=\log D_{AB}$ and the 3-phase currents I_i are balanced, thus displaced 120 electrical degrees apart, say, $I_B=I_A(-0.5+j0.866)$ and $I_C=I_A(-0.5-j0.866)$. Substituting accordngly in equation 8 yields

$$L_A=-2(\log D_{AA}- \log D_{AB}) \quad (9)$$

Now by equation 27 of reference 15, the self-geometric mean distance D_{AA} of conductor A is given by

$$(S_1+S_2)^2 \log D_{AA}=\sum_{i=1}^{2} \sum_{j=1}^{2} S_iS_j \log D_{ij} \quad (10)$$

whence inasmuch as $S_1=S_2$

$$\log D_{AA}=(1/2)[\log D_{11}+\log D_{12}] \quad (11)$$

Again, from equation 27 of reference 15, the mutual geometric mean distance D_{AB} between conductors A and B is given by

$$(S_1+S_2)(S_3+S_4) \log D_{AB} =\sum_{i=1}^{2} \sum_{j=3}^{4} S_iS_j \, 1n \, D_{ij} \quad (12)$$

whence, inasmuch as $S_1=S_2=S_3=S_4$ and $D_{13}=D_{24}$

$$\log D_{AB}=(1/4)[2 \log D_{13}- \log D_{14}+ \log D_{23}] \quad (13)$$

Substituting from equations 11 and 13 in equation 9 as indicated yields

$$L_A=-[\log D_{11}+ \log D_{12}- \log D_{13}-(1/2)\times \log D_{14}-(1/2) \log D_{23}] \quad (14)$$

The five required geometric mean distances D_{ij} are derived in the Appendix. Substituting therefrom in equation 14 as indicated, collecting corresponding terms, and defining K by $K=(C/\sqrt{3}a)$ yields the following general equation for the inductance $L=L_A=L_B=L_C$ associated with each of the three conductors of the bus

$$L=(1/8)[12K^2\ln K+(1-6K-3K^2)\times \ln(1+3K^2)-(2+4.5K+1.5K^2)+ \ln(1+3K+3K^2)+3(2+2K-K^2)\ln(1+K+K^2)+6(1+2K+K^2)\ln(1+K)]+\sqrt{3}/4\{(1-2K) \tan^{-1} [\sqrt{3}(1-K)/(1+3K)]+3(2K+K^2) \tan^{-1} [\sqrt{3}(1-K)/(3+K)]+(1+4K+3K^2) \tan^{-1} [\sqrt{3}K/(2+3K)]+(1+2K) \tan^{-1} [1/\sqrt{3}\times (1+2K)]+(4K-3K^2) \tan^{-1}\times (1/\sqrt{3}K)\}+(3/8)(3K+K^2)\ln\times a+(3/4)(0.5+2K+K^2)\ln 3-\pi/4\sqrt{3} \quad (15)$$

Corroboration of Equation 14

A check on the validity of the general equation 15 is to be obtained as follows. If the spacing c between the angles is set equal to zero, the resulting configuration is that of Fig. 3. In such case, $D_{12}=D_{13}=D_{14}$ and $D_{23}=D_{11}$. Substituting accordingly in equation 14 yields

$$\log D_{AA}=-(1/2)[\log D_{11}-\log D_{12}] \quad (16)$$

Substituting for the geometric mean distances from the Appendix as indicated yields

$$L=(\pi/8\sqrt{3})+(3/8)\ln 3 \quad (17)$$

Corroboratively, substituting $c=0$ in equation 15 yields equation 17.

Determination of Electromagnetic Forces

By assuming an origin of rectangular co-ordinates at the centroid of cross section of the bus with x axis directed, say, parallel to the line segments S_1 and S_6 of Fig. 2, general expressions for the electromagnetic force on any conductor by means of equation 6 could be calculated. However, such equations would prove so lengthy and complicated in form as to render them more or less impractical for numerical computation. A rather simple and sufficiently satisfactory procedure for calculating the electromagnetic force on a conductor is afforded by the following considerations.

In virtue of the geometrical symmetry and of the close spacing between adjacent flanges, the force exerted on a conductor will be primarily caused by the vector sum of the forces exerted on its two flanges by the electromagnetic forces produced on these flanges by the currents in the immediately adjacent flanges of the other conductors. Obviously, the effective mutual electromagnetic force between adjacent flanges will be normal to the flanges; and these flanges of themselves have the shapes of strap conductors. These observations thus suggest that the force acting on the conductor be approximated by finding the forces exerted on its two flanges by the immediately adjacent flanges through treating a pair of adjacent flanges as a pair of parallel-faced strap conductors, each of which carries 1/2 of the corresponding conductor current. But, in such a case, the desired flange pair force can be obtained directly from the well-known curves for the electromagnetic force between two parallel-faced strap conductors given by Dwight[30] and reproduced in reference 2, page 37.

Units

In that it is yet customary in actual bus design to express distances in terms of feet, force in terms of pounds, and so forth, and in that the forms of and subsequent calculations with the general equations are somewhat simplified if centimeter-gram-second rather than meter-kilogram-second units are used, experience is to the effect that it is convenient to advance all equations in terms of centimeter-gram-second units and, subsequently, convert to the various units used in actual design by multiplication with appropriate conversion factors.

Accordingly, with regard to the physical quantities yielded by the equations mentioned to this point, all units, aside from certain permissible and desirable exceptions described in subsequent paragraphs, are in the absolute system: linear dimensions in centimeters, current in abamperes, inductance in abhenrys per centimeter of bus length, energy in ergs per centimeter of bus length, and force in dynes per centimeter of bus length. Multiplying abhenrys per centimeter by 3.048×10^{-5} yields millihenrys per foot of bus length; multiplying by 11.48×10^{-6} yields 60-cycle reactance in ohms per foot of bus length.

In equations for the total inductance associated with the kth conductor, the dimensions of the cross section of the bus can be expressed in any like unit of length. This is to be seen as follows: In equation 1, changing the unit in which D_{ik} is expressed is equivalent to multiplying D_{ik} by a constant, say, s. In turn, the right-hand member of the resulting equation can be written as the sum of the right-hand member of equation 1 plus the term

$$-2\sum_{i=1}^{r} (I_i/I_k) \log s$$

But this term equals zero because of the relationship expressed by equation 4.

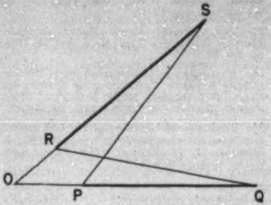


Fig. 4 (left). General disposition of two line segments

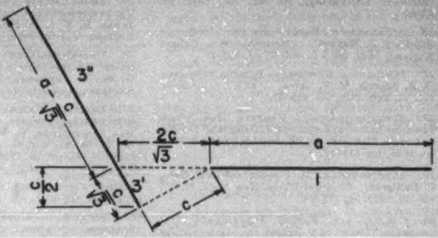


Fig. 5 (right). Cross section for calculation of typical geometric mean distance

Hence, regardless of the unit chosen, the actual inductance is given by equation 1 wherein the dimensions of the cross section of a particular bus can be expressed in terms of the unit of length best suited to ease of calculation for the bus under consideration.

Again, in computing forces, it may prove inconvenient to express the bus dimensions in centimeters since design values are commonly expressed in inches or feet. However, it is easily and similarly established that the dimensions can be expressed in terms of any convenient unit of length, to the end that the computed forces are expressed in terms of dynes per same unit of bus length. Subsequent multiplication by 2.248×10^{-6} yields force in pounds per chosen unit of length.

Thus, to summarize, the desired inductances and forces can be calculated from the derived general equations; therein linear dimensions can be expressed in units of inches or feet (if so desired); and the resulting computed values can be converted directly to the mentioned convenient working units by use of the stated conversion factors.

An Illustrative Example

To illustrate use of equation 15, consider a bus of the form of Fig. 2 having the following typical dimensions: Thickness $t = 3/16$ inch; spacing $d = 1/4$ inch; and width W along the inside of the angle $= 4$ inches. The following values result: $c = (1/4) + (3/16) = 7/16 = 0.4375$ inch; $a = 4 - (1/2\sqrt{3})(3/16) = 3.94587341$; $K = c/\sqrt{3}a = 0.0640139$; and $K^2 = 0.0040978$. Substituting these values in equation 15, obtaining values of the logarithms and arc tangents from the 15-place Works Project Administration tables, and effecting all computations on an electric calculating machine yields $L = 1.054$ abhenrys per centimeter of bus length. Multiplying this value by 3.048×10^{-6} yields $L = 3.21 \times 10^{-6}$ millihenrys per foot of bus length.

Corroboratively, check of this value against corresponding values of several busses of comparable current rating comprised of more conventionally arranged strap and channel conductors reveals that the just computed value for the nested 120-degree-angle bus to be about one-third to one-half as great as for these other conductors; hence the reactive voltage drop will be in the same proportion. This comparison reveals the considerable efficacy of the nested 120-degree-angle bus for low-voltage heavy-current duty where the reactive voltage drop must be minimized to obtain a specified permissible limit of voltage regulation, and also is in agreement with similar results advanced in reference 4, where the inductive reactance of a 3-phase 5,000-ampere finned 120-degree-angle bus had about 40 per cent of the inductive reactance of a similarly rated bus comprised of 7-inch double-channel conductors coplanarly spaced on 12-inch centers.

Appendix. Calculation of the Five Geometric Mean Distances

Calculation of the five geometric mean distances indicated in equation 14 is to be typified as follows. From page 327 of reference 16, the geometric mean distance D between the line segments RS and PQ of Fig. 4 is given by

$$\begin{aligned}(PQ)(RS)[2 \log D+3] = {} & \log (PR)\times \\ & [2(OP)(OR) \sin^2\theta-(PR)^2 \cos\theta]+ \\ & \log (QS)[2(OQ)(OS) \sin^2\theta-(QS)^2 \\ & \cos\theta]- \log (PS)[2(OP)(OS) \sin^2\theta- \\ & (PS)^2 \cos\theta]- \log (QR)[2(OQ)(OR) \\ & \sin^2\theta-(QR)^2 \cos\theta]- \sin\theta[(OP)^2 \\ & \text{angle } SPR-(OQ)^2 \text{ angle } SQR+(OR)^2 \\ & \text{angle } PRS-(OS)^2 \text{ angle } PSQ] \qquad (18)\end{aligned}$$

where θ designates the angle at the vertex O of Fig. 4, and the distances from one point on the line to another are given in parentheses.

To calculate D_{13} by use of this equation is shown as follows. With respect to Fig. 5, by equation 17 of reference 15, it is seen that

$$a^2 \log D_{13} = a(a-d/\sqrt{3}) \log D_{13''} + (ac/\sqrt{3}) \log D_{13'} \qquad (19)$$

On comparing Figs. 4 and 5, by aid of equation 18 the two geometric mean distances of $\log D_{13''}$ and $\log D_{13'}$ in equation 19 can be calculated. The computation is tedious but straightforward. In similar fashion, the other four geometric mean distances of equation 14 can be calculated; or else, as for $\log D_{11}$ and $\log D_{33}$, they can be written down from equations given in reference 16. For complete details see reference 31.

References

1. COLE INDUSTRIES INCORPORATED. *U. S. Patent No. 2,400,223*, May, 1946.

2. CHASE ELECTRICAL HANDBOOK (book). Chase Brass and Copper Company, Inc., Waterbury, Conn., 2nd ed., 1947.

3. FEATURES OF A NEW A-C BUS DESIGN, R. N. Wagner. *Aluminum Company of America*, Pittsburgh, Pa., 1948.

4. ALUMINUM CONDUCTORS USED EXCLUSIVELY THROUGHOUT NEW ROLLING MILL. *Electrical World*, New York, N. Y., vol. 129, April 24, 1948, pp. 81–84. (Comprises much of reference 3.)

5. ALUMINUM CHARACTERISTICS FOR BUS APPLICATION, L. T. Guess. *Electrical World*, New York, N. Y., vol. 131, April 23, 1949, pp. 90–91.

6. ALCOA ALUMINUM BUS CONDUCTORS (book). Aluminum Company of America, Pittsburgh, Pa., 1952.

7. THE DESIGN OF BUS-BAR INDUSTRIAL DISTRIBUTION SYSTEMS: AN EPITOMIZATION OF AVAILABLE DATA, T. J. Higgins. *AIEE Transactions*, vol. 65, 1946, pp. 328–36.

8. TEMPERATURE RISE OF BUS BARS, H. B. Dwight, G. W. Andrew, H. W. Tileston. *General Electric Review*, Schenectady, N. Y., vol. 43, 1940, pp. 213–16.

9. THE CALCULATION OF THE SHORT-TIME RATINGS OF BARE ELECTRICAL CONDUCTORS. L. H. R. Everitt. *Journal*, Institution of Electrical Engineers, London, England, vol. 93, 1945, pp. 380–87.

10. REACTANCE OF SQUARE TUBULAR BUS BARS, H. B. Dwight, T. K. Wang. *AIEE Transactions*, vol. 57, 1938, pp. 762–66.

11. FORMULAS FOR THE INDUCTANCE OF COAXIAL BUSSES COMPRISED OF SQUARE TUBULAR CONDUCTORS, H. P. Messinger, T. J. Higgins. *AIEE Transactions (Electrical Engineering)*, vol. 65, June 1946, pp. 328–36.

12. THE INDUCTANCE OF LINEAR CONDUCTORS OF RECTANGULAR SECTION, A. H. M. Arnold. *Journal*, Institution of Electrical Engineers, London, England, vol. 70, 1931–32, pp. 579–86.

13. EQUATIONS FOR THE INDUCTANCE AND SHORT-CIRCUIT FORCES OF BUSSES COMPRISED OF DOUBLE-CHANNEL CONDUCTORS, C. M. Siegel, T. J. Higgins. *AIEE Transactions*, vol. 71, pt. III, Jan. 1952, pp. 522–31.

14. FORMULAS FOR THE GEOMETRIC MEAN DISTANCES OF RECTANGULAR AREAS AND OF LINE SEGMENTS, T. J. Higgins. *Journal of Applied Physics*, New York, N. Y., vol. 14, 1943, pp. 188–95.

15. FORMULAS FOR THE CALCULATION OF THE INDUCTANCE OF LINEAR CONDUCTORS OF STRUCTURAL SHAPE, T. J. Higgins. *AIEE Transactions*

(*Electrical Engineering*), vol. 62, Feb. 1943, pp. 53–57.

16. A TREATISE ON ELECTRICITY AND MAGNETISM (book), **J. C. Maxwell.** Clarendon Press, Oxford, England, 1892, 3rd ed., vol. 2, pp. 326–27.

17. MECHANICAL STRESSES IN BUS BAR SUPPORTS, **O. R. Schurig, M. F. Sayre.** *AIEE Transactions*, vol. 44, Feb. 1925, pp. 213–37.

18. ELEKTRODYNAMISKA KRAFTER HOS STROM FORANDE STÄLLVERKSSKENOR, **F. Dahlgren.** *Teknisk Tidskrift Elektroteknik*, Stockholm, Sweden, vol. 55, 1925, pp. 117–21.

19. ELECTRODYNAMICAL FORCES IN BUS BARS, **F. Dahlgren.** *Electrical World*, New York, N. Y., vol. 86, 1925, pp. 473–74.

20. CALCUL DES FORCES INTÉRIEURES DÉVELOPÉES DANS DES SYSTEMES MECHANIQUES SOUMÉS A DES ÉFFORTS ELECTRODYNAMIQUES, **A. Eigermann.** *Rêvue Génerale de l'Électricité*, Paris, France, vol. 35, 1934, pp. 677–88.

21. SHORT-CIRCUIT FORCES ON BUS BARS AND THEIR SUPPORTS IN CELLULAR SWITCHGEAR STRUCTURES, **E. E. I. Pilcher.** *World Power*, London, England, vol. 24, 1935, pp. 118–23 and 231–4.

22. UNTERSUCHUNGEN ÜBER ELEKTRODYNAMISCHE BEANSPRUCHUNG, **M. F. Aebi.** *Schweizer Archiv für Angewandte Wissenschaft und Technik*, Solothurn, Switzerland, vol. 7, 1941, pp. 185–95.

23. PRACTICAL CALCULATION OF SHORT-CIRCUIT STRESSES IN SUPPORTS FOR STRAIGHT PARALLEL BAR CONDUCTORS, **O. R. Schurig, C. W. Frick, M. F. Sayre.** *General Electric Review*, Schenectady, N. Y., vol. 29, 1926, pp. 534–44.

24. STRESSES IN BUS SUPPORTS, **R. Tanberg.** *Electric Journal*, East Pittsburgh, Pa., vol. 24, 1927, pp. 517–25.

25. SHORT-CUT METHODS OF CALCULATING STRESSES IN BUS STRUCTURES, **W. Specht.** *General Electric Review*, Schenectady, N. Y., vol. 31, 1928, pp. 431–18.

26. SHORT-CIRCUIT STRESS CHARTS, **W. M. Edson.** *Electrical World*, New York, N. Y., vol. 94, 1929, pp. 112–15.

27. BUS STRESSES UNDER SHORT CIRCUITS. *Electrical World*, New York, N. Y., vol. 92, 1928, pp. 564–65.

28. FORMULAS FOR CALCULATING SHORT-CIRCUIT FORCES BETWEEN CONDUCTORS OF STRUCTURAL SHAPE, **T. J. Higgins.** *AIEE Transactions (Electrical Engineering)*, vol. 62, Oct. 1943, pp. 659–63.

29. FORMULAS FOR CALCULATING THE INDUCTANCE OF CHANNELS LOCATED BACK TO BACK, **W. G. Schwantz, T. J. Higgins.** *AIEE Transactions*, vol. 65, 1946, pp. 893–99.

30. REPULSION BETWEEN STRAP CONDUCTORS, **H. B. Dwight.** *Electrical World*, New York, N. Y., vol. 70, 1917, pp. 522–24.

31. EQUATIONS FOR THE DESIGN OF THREE-PHASE BUSSES COMPRISED OF 120°-ANGLES, **J. W. Maier.** *M.S. Thesis*, University of Wisconsin, Madison, Wis., 1951.

◆

Discussion

L. P. Lanigan (Chase Brass and Copper Company, Waterbury, Conn.): One of the important considerations of most a-c busses is of course the reactance voltage drop. When the bus is of considerable length and the current is relatively high, careful attention must be given to this factor. There are various methods of designing an electric bus so as to maintain the reactance voltage drop at a low value.

The 3-phase bus design consisting of 120-degree-angle conductors described in the paper provides an effective method of minimizing the reactance voltage drop. Although this is the most important advantage of such a bus, it has other desirable characteristics as pointed out by the author. The bending of the conductors to a 120-degree angle provides mechanical strength to withstand short-circuit forces. Its compactness minimizes space while the flat surfaces permit the connecting of branch circuits.

The copper suppliers can furnish the straight lengths of flat rectangular bus bars to the bus fabricating companies. The latter can then bend them into the required shapes and complete the assembly of the bus. It would appear that consideration could be given to this type of bus when the bus is long and the current is quite high so as to maintain the reactance voltage drop at a low value.

Donald H. Sandell (Aluminum Company of America, Pittsburgh, Pa.): We believe this is a worth-while paper which makes a definite contribution to the store of knowledge that bus designers can use in their work.

There are many installations in our company where sector bus (as we call it) is used and we have knowledge of some outside of our company. In addition to generating station and industrial use, this bus is used to distribute power to load centers in large commercial buildings. Our test data deviate in some minor respects from the conclusions in the paper, but are not inconsistent with the general relationships given in the paper. For instance, our tests indicate that the nonuniform current distribution in the conductors is significant as evidenced by the ratio of a-c to d-c resistance of from about 12 to over 20 per cent.

Temperature rise measurements also indicate significant nonuniform current density. It must be admitted, however, that these tests were conducted on practical sector bus shapes and not on the ideal uniform shapes as shown in the paper.

C. M. Siegel (University of Virginia, Charlottesville, Va.): In view of the increasing importance of busses comprised of 120-degree-angle conductors, this paper should be of appreciable value to engineers concerned with the design of busses having low reactance and considerable mechanical rigidity. The authors have given a clear development, leading to results that are readily applicable to the rational design of a bus having the subjective configuration. The utility of the results is enhanced by the authors' discussion and references which justify the assumptions and approximations that are necessary for a tractable analysis.

Equation 15 of the paper is well suited for simplification by expansion in a power series. If $K \ll 1$, then power series expansions of the terms involving the logarithms and arctangents, about $K=0$, are rapidly convergent. Since in practice the dimension a, Fig. 2, will often be much larger than the dimension c, thus giving $K \ll 1$, such expansions followed by collection of terms involving like powers of K will reduce equation 15 to a very simple form.

Expansions of the terms involving the logarithms and arctangents are facilitated by the following relations[1]

$$\log (1+x) = x - x^2/2 + x^3/3 - \ldots \quad (x^2<1 \text{ and } x=1)$$

$$\tan^{-1} [(m+nx)/(p+qx)] = \tan^{-1}(m/p) + x(np-mq)/(p^2+m^2) + (x^2/2!)2(mq-np)(mn+pq)/(p^2+m^2)^2 + \ldots$$

Applying these to the various terms (excepting the first logarithmic term) of equation 15, multiplying polynomials where indicated, and collecting coefficients of like powers of K yields

$$L = (\pi/8\sqrt{3}) + 0.375 \log 3 + 1.5K^2 \log K + K[(2\pi/\sqrt{3}) + 1.5 \log 3 + 1.125 \log a] + K^2[-0.5625 - (3\pi/4) + 0.375 \log a + 0.75 \log 3] + \ldots \quad (20)$$

A check on the validity of this equation is effected by noting that the first two terms (to which the right side reduces when $K=0$) agree with equation 17 of the paper. This leads us to note that the inductance is expressed in equation 20 in terms of the value for coinciding flanges ($c=0$, $K=0$) plus departures from this value due to separation ($c \neq 0$, $K \neq 0$) between the flanges.

Simplifying the combining the constant quantities in equation 20 yields, finally

$$L = 0.63871 + 1.5K^2 \log K + K(5.2755 + 1.125 \log a) + K^2(-1.0989 + 0.375 \log a) + \ldots$$

Applying this to the illustrative example of the authors' paper and effecting computations with a slide rule, we obtain

$$L = 0.6387 - 0.0169 + 0.4365 - 0.0024 + \ldots = 1.056 \text{ abhenrys per centimeter of bus length}$$

This agrees with the value in the paper.

We see, in this example, that neglecting the value -0.0024 associated with the last term increases the error by less than 0.25 per cent. Hence it follows that, for busses having $K \ll 1$, sufficient accuracy for engineering purposes is obtained by neglecting terms involving K to more than the first power (excepting the log K term).

REFERENCE

1. TABLES OF INTEGRALS AND OTHER MATHEMATICAL DATA, **H. B. Dwight.** The Macmillan Company, New York, N. Y., 1947, equations 601 and 38.

J. W. Maier and **T. J. Higgins**: Mr. Lanigan has very aptly summarized the essential electrical and mechanical advantages of the 3-phase 120-degree-angle bus, as viewed by a practicing bus designer. His remarks both substantiate and compliment similar statements by the authors at various points in their paper.

The authors are more appreciative of Mr. Sandell's complimentary appraisal of the general merits of their paper. Additively, they are pleased to have Mr. Sandell's statement and reference as to successful use of the 3-phase 120-degree-angle bus for power supply to load centers of commercial buildings, a possible and important use not specifically remarked in their paper.

Mr. Sandell's remark that tests reveal that skin and proximity effects produce an a-c resistance somewhat greater than the d-c resistance, with a resulting slight nonuniform current-density distribution as evidenced by temperature rise measurements, is in accordance with a similar statement in the beginning of our paper. Further, the values he cites confirm other of our remarks: that the actual increase of resistance will not be large; and that in view of such slight increase (which indicates a corresponding relatively small shift from uniform current distribution), the values of inductance calculated on the assumption of uniform current distribution will yet be substantially the same as for uniform distribution.

It is of interest to note that these facts for the rather complicated 120-degree-angle geometry are in accord with similar variations in conductors of circular or annular cross section whereof the skin-effect changes of inductance and resistance with increase of frequency can be calculated exactly. Moreover, as inductance decreases with increases of frequency, a value calculated on the assumption of uniform current distribution will be slightly higher than the actual value; whence the calculated value will be on the optimum side, as regards use in design.

The authors are indebted to Dr. Siegel for an admirable supplement to the analysis of their paper. By use of series expansions in terms of small parametric ratios of bus geometry, a technique successfully employed earlier in his work with one of the authors on double-channel busses,[1,2] Dr. Siegel has advanced equations which enable very rapid and accurate slide-rule calculation of the desired inductance.

In conclusion, we extend our most sincere thanks for the valuable substantiative and illuminative comment advanced by the three discussers of our paper.

References

1. See reference 13 of the paper.
2. Equations for Determining Current Distribution Among the Conductors of Busses Comprised of Double-Channel Conductors, Clifford M. Siegel, Thomas J. Higgins. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 489–92.

An Elliptical Polarization Synthesizer

G. H. FRIEDMAN
ASSOCIATE MEMBER AIEE

Synopsis: An electronic instrument is described which automatically determines the resultant polarization of two superimposed elliptically polarized waves when the polarization of the individual components is known. In addition, a study of the effect of ground pulse contamination on polarization data obtained by sounding the ionosphere is presented. The study serves to demonstrate the usefulness of the instrument.

IN THE investigation of the properties of the ionosphere through the study of the observed polarization of ionospherically reflected low-frequency radio waves, it often becomes necessary to compute the resultant polarization of two component waves. More often, the two components are related magneto-ionically, but occasionally problems arise in which the components are more generally related or not related at all. Methods are available for computing this resultant either analytically or graphically when the polarization, relative amplitude, and phase difference of the two component waves are known.[1,2] Although these methods are not complex, they tend to become lengthy and tedious whenever many computations are necessary and, very often, the physical significance of the results tends to become obscured.

In conjunction with a program of ionospheric research conducted at the Ionosphere Research Laboratory, an instrument has been built which performs the computation for obtaining the resultant polarization of the two component waves automatically.[2] This resultant is graphically represented on the screen of a cathode-ray oscillograph. The parameters of the ellipse, angle of the tilt of the major axis, axial ratio and rotational sense may be observed and measured directly on the oscillograph screen. The advantages of using such an instrument are twofold. In addition to the saving of time and labor which is consumed in computational methods, it is possible for the operator to reproduce the observed behavior of an ionospherically reflected echo and thereby study, in a semi-quantitative way, the variation of the individual component waves. This leads to a fuller understanding of the physical processes involved.

The Elliptical Polarization Pattern

That property of electromagnetic radiation known as polarization, the process by which the variations of the electric and magnetic intensity of a wave are confined to a definite pattern, is familiar enough so that very little detailed discussion of the concept need be given here. It will be sufficient to recall that the polarization pattern is a graphical representation of the locus of either the electric or magnetic intensity vector of the polarized wave. For single-frequency radiation that locus is, in general, an ellipse. It is well known that the vector describing the ellipse may be resolved into sinusoidally varying components in space quadrature.

A set of rectangular co-ordinate axes, x and y, and a typical polarization ellipse are shown in Fig. 1. The value of ρ_1 represents the ratio of the amplitudes of the sinusoidal space quadrature components along the x and y axes. A second set of rectangular co-ordinate axes can be defined to coincide with the major and minor axes of the ellipse. The angle of tilt of the ellipse ψ is defined as the angle between the two sets of co-ordinate axes. The angle θ is defined as the angle whose tangent is the ratio of the length of the minor axis to the length of the major axis. The quantities ψ, θ, and the direction of rotation of the vector serve to define completely the ellipse, neglecting absolute magnitudes. It can be shown that these three parameters may be simply written in terms of the ratio ρ_1 and the phase difference between the sinusoidal components along the x and y axes.

When a sinusoidal voltage is applied to the vertical deflection plates of a cathode-ray tube and another sinusoidal voltage of the same frequency is applied to the horizontal deflection plates, an elliptical Lissajous pattern will, in general, be traced out on the cathode-ray tube screen. If desired, this Lissajous pattern can be caused to be an exact representation of the elliptical polarization pattern of some particular electromagnetic wave. To do this, the ratio of the amplitudes and the phase difference of the two voltages

---

Paper 55-155, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted May 29, 1953; made available for printing December 15, 1954.

G. H. Friedman is with The Franklin Institute Laboratories for Research and Development, Philadelphia, Pa.

The research reported in this paper has been sponsored by the Geophysics Research Division of the Air Force Cambridge Research Center under contract *AF19(122)-44*.

The work reported in this paper constitutes a portion of a thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Electrical Engineering at the Pennsylvania State College. Grateful acknowledgment of the support of an Ordnance Research Laboratory Fellowship should also be made. Thanks are due to Prof. Harold Nearhoof for his assistance in interpreting the experimental results and to Carl Berkley of Allen B. DuMont Laboratories for his suggestion of the pulse modulation scheme for indicating rotation.

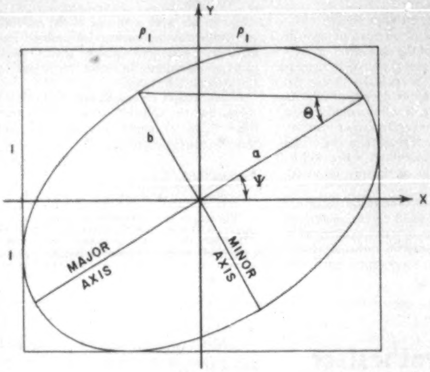

Y
ρ1
ρ1
θ
b
a
Ψ
X
MAJOR AXIS
MINOR AXIS

Fig. 1. Polarization ellipse showing co-ordinate system and axes

may be set equal to the same parameters of the electric or magnetic intensity space quadrature components.

Finding the resultant polarization pattern of two superimposed waves when the polarization of each is known is essentially a process in which the corresponding space quadrature components of the appropriate vector for each wave are added by means of the theorems of trigonometry. Here again, to produce a Lissajous pattern which represents the polarization of the two superimposed waves, it is merely necessary to impress voltages upon the deflections plates of the cathode-ray tube which correspond to the sums of the wave quadrature components.

The General System

In a relatively straightforward manner, it is possible to construct an electronic instrument which will provide the four sinusoidal voltages necessary to represent the space quadrature components of the two polarized waves it is desired to study. This is done by suitable use of attenuators, phase shifters, switches, and amplifiers. The corresponding pairs of these voltages may then be added and the resultant signals used to energize a cathode-ray oscillograph. The functional operation of such an instrument is shown in Fig. 2.

The attenuator A and the phase shifter α, shown in Fig. 2, provide a means for inserting quantities which account for the relative amplitude and total phase difference between the two waves whose resultant polarization pattern is to be found. The total channel switch enables the operator to apply these quantities to either of the two ellipses which are being formed. The attenuators ρ_1 and ρ_2 are provided in one line of each of the two ellipse forming channels so that the proper value for the amplitude ratio of the two voltages used to form each ellipse may be inserted. By proper setting of the attenuator axis switches, these attenuators may be used in either of the two lines of its channel.

This feature is necessary since, to be able to form an ellipse with any given axial ratio and angle of tilt, the amplitude of one of the quadrature components must be from zero to infinity or, as is more practical, the magnitude of each of the quadrature components must be varied with respect to each other. The channel selector switch provides a means for viewing each of the component ellipses which are formed or their resultant, as desired.

To define completely the polarization pattern, it is also necessary to specify the direction of rotation of the vector or of the corresponding cathode-ray tube trace. On the instrument that was built, this determination was made possible by a circuit which shifts the vertical deflection

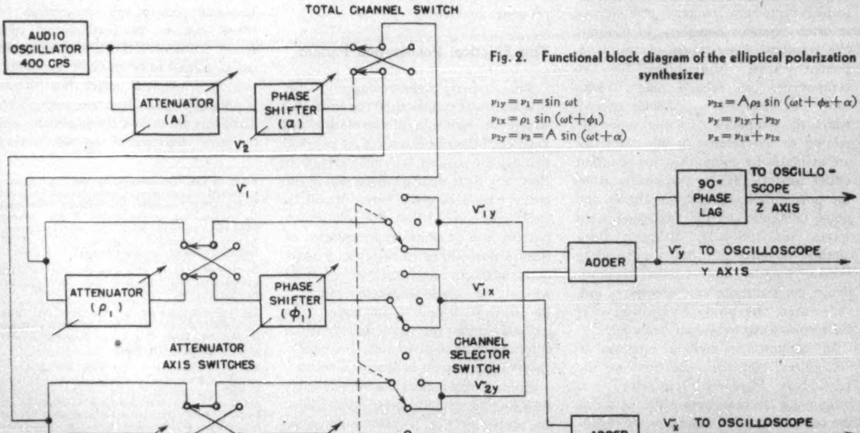

TOTAL CHANNEL SWITCH
AUDIO OSCILLATOR 400 CPS
ATTENUATOR (A)
PHASE SHIFTER (α)
V2
V1
ATTENUATOR (ρ1)
PHASE SHIFTER (φ1)
ATTENUATOR AXIS SWITCHES
ATTENUATOR (ρ2)
PHASE SHIFTER (φ2)
V1y
V1x
CHANNEL SELECTOR SWITCH
V2y
V2x
ADDER
Vy TO OSCILLOSCOPE
Y AXIS
90° PHASE LAG
TO OSCILLO-SCOPE
Z AXIS
ADDER
Vx TO OSCILLOSCOPE
X AXIS

Fig. 2. Functional block diagram of the elliptical polarization synthesizer

$v_{1y} = v_1 = \sin \omega t$

$v_{1x} = \rho_1 \sin (\omega t + \phi_1)$

$v_{2y} = v_2 = A \sin (\omega t + \alpha)$

$v_{2x} = A\rho_2 \sin (\omega t + \phi_2 + \alpha)$

$v_y = v_{1y} + v_{2y}$

$v_x = v_{1x} + v_{2x}$

1, 74
1
55

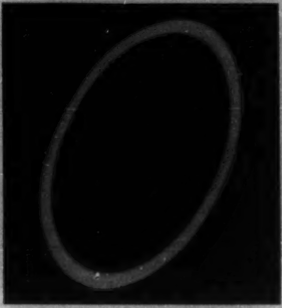


Fig. 3. Typical polarization pattern

voltage 90 degrees lagging in phase. This phase-shifted voltage is then coupled to the intensity modulation amplifier terminals of the oscillograph. When this is properly done, half of the elliptical pattern will be blanked out. A study of Lissajous patterns will show that, for a pattern produced by a trace rotating in the counterclockwise, or left-handed, direction, only the left-half portion of the pattern will remain visible. Although this method for determining the direction of rotation serves satisfactorily, a somewhat simpler scheme has been suggested to the author. By modulating the intensity of the trace with pulses possessing a sharp rise and exponential decay, a bright spot with a tail will be produced. The direction in which the tail will point will indicate the direction of rotation.

A typical polarization pattern is shown in Fig. 3 and a modulated pattern in Fig. 4.

Circuit Details

Fig. 5 is a schematic diagram of the more important circuits of the elliptical polarization synthesizer which was built. Standard resistance-capacitance coupled amplifiers and cathode followers are used throughout. The attenuators used are simple potentiometer voltage dividers. Of particular interest, perhaps, are the continuously variable linear phase-shifter circuits.[3] These employ resolver-type synchros or, simply, resolvers. Resolvers possess a rotor with a single winding and two stator windings in space quadrature. When the rotor winding is energized by a

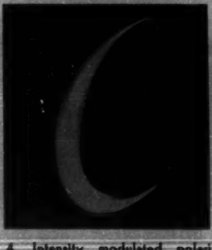


Fig. 4. Intensity modulated polarization pattern showing left-handed rotation

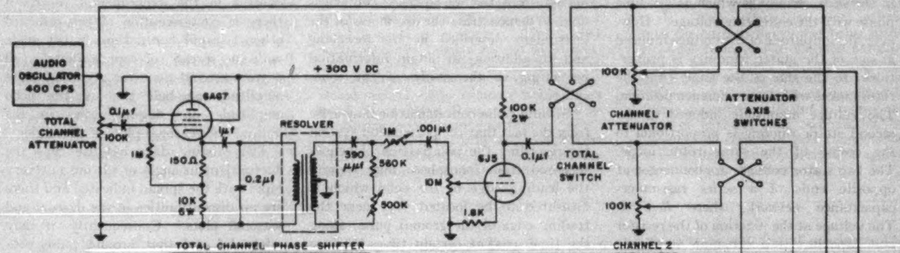

AUDIO OSCILLATOR 400 CPS
TOTAL CHANNEL ATTENUATOR
100K
0.1µf
6AG7
+300 V DC
1M
150Ω
10K 6W
1µf
0.2µf
RESOLVER
390 µµf
1M
.001µf
560K
500K
10M
6J5
1.8K
100K 2W
0.1µf
TOTAL CHANNEL SWITCH
TOTAL CHANNEL PHASE SHIFTER
100K
CHANNEL 1 ATTENUATOR
ATTENUATOR AXIS SWITCHES
100K
CHANNEL 2 ATTENUATOR
+300V DC
6AG7
0.1µf
1M
150Ω 1W
10K 6W
1µf
0.2µf
RESOLVER
390 µµf
1M
560K
500K
CHANNEL 1 – PHASE SHIFTER
2M
300K
10M
10M
2M
300K
10M
10M
CHANNEL SELECTOR
6AG7
0.1µf
1M
150Ω 1W
10K 6W
1µf
0.2µf
RESOLVER
390
1M
560K
500K
CHANNEL 2 – PHASZ SHIFTER
510K 2W
INTENSITY MODULATION OUTPUT
700µµf
.001µf
10M
10M
6J5
510K 2W
VERTICAL DEFLECTION OUTPUT
.01µf
500K
.001µf
10M
½ 6SN7
510K 2W
HORIZONTAL DEFLECTION OUTPUT
.01µf
500K
.001µf
10M
½ 6SN7

Fig. 5. Elliptical polarization synthesizer—schematic diagram

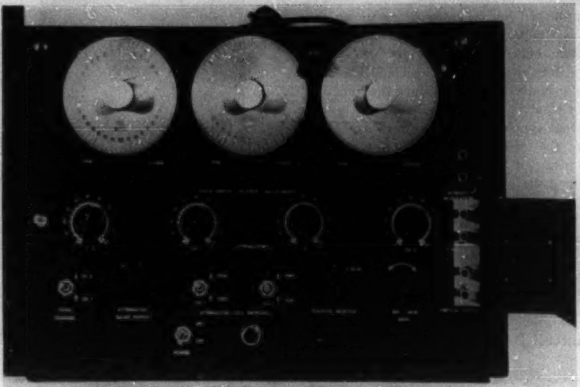


Fig. 6. Elliptical polarization synthesizer—front panel view

sinusoidal voltage, voltages are induced in the stator windings which are in time phase with the energizing voltage. However, the amplitude of the voltage induced in one of the stator windings is proportional to the sine of the angle that the rotor makes with some reference position. The voltage amplitude induced in the second stator winding is proportional to the cosine of the same rotor angle. The two stator voltages are connected at opposite ends of a series resistance-capacitance network where $R\omega C=1$. The voltage at the junction of the resistor and capacitor is a constant amplitude signal whose phase varies linearly and continuously with the angular position of the resolver rotor.

Experimental Results

A study was conducted to determine the possible effect of ground pulse contamination on the observed polarization of a 150-kc-per-second pulsed signal reflected from the ionosphere. This was done to demonstrate the usefulness of the instrument described in the foregoing and, in addition, to obtain information pertaining to the study of the ionosphere.

Ground pulse contamination may arise from the fact that the polarimeter system for recording the polarization of echoes returned from ionosphere may require the leading edge of the echo which is measured to be located very near the trailing edge of the ground pulse, along the time axis, at certain times.[4] Since the recording equipment chooses a particular portion of the time axis by a suitable gate, it is conceivable that a portion of the ground pulse in contained within the gate.

This problem was set up on the elliptical polarization synthesizer, Fig. 6, in such a way that various magnitudes of ground pulse could be added to the returned echo polarization with different phases and relative amplitudes between the two. The echo polarization chosen for the study was typical of what might be expected under conditions where pulse overlapping is possible.

The problem was to determine whether or not the addition of various amounts of ground pulse could produce the phenomenon locally called flip-overs. This is the rotation of the major axis of the net ellipse into the second and fourth quadrants. Flip-overs cannot be explained by simple theory for the frequency of operation (150 kc per second) and geographical location (State College, Pa.) if a pure magnetic-ionic wave is being returned.

By varying the magnitude of the ground pulse polarization pattern and by varying the phase difference between the ground pulse and the echo signal, the effect on the resultant polarization was determined. Curves of this variation are plotted in Figs. 7 and 8. The curves indicate two things that should be evident in the experimental results if there is contamination of the returned echo: 1. rapid fluctuations in tilt angle with a spread of approximately 40 degrees around noon for a ground pulse amplitude one-half that of the echo amplitude; 2. discontinuities in the diurnal and seasonal plots of tilt angle.

The existing data indicate that the fluctuations in angle of tilt are not very rapid with the spread indicated and there are no discontinuities in the diurnal and seasonal plots.[5] Consequently, it may safely be said that ground pulse contamination does not cause the so-called flip-overs to the extent recorded experimentally.

The particular polarization pattern chosen to represent the echo is one possessing an angle of tilt ψ equal to 74 degrees and axial ratio represented by the angle θ equal to 30 degrees, and which possesses a positive sense of rotation. The ground pulse polarization pattern, corresponding to the actual observed

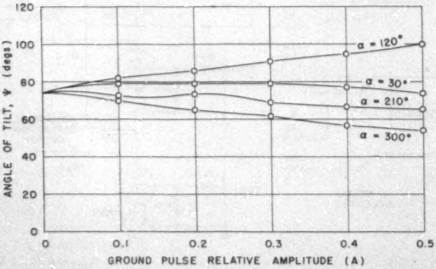


Fig. 7. Variation of angle of tilt ψ for contaminated echo

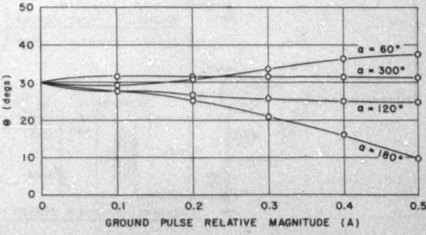


Fig. 8. Variation for Θ for contaminated echo

pattern, is one with an angle of tilt ψ of 11 degrees, θ equal to 30 degrees, and positive rotation. Data were taken using the elliptical polarization synthesizer in the following manner. With the channel 1 and 2 controls set to produce the echo and ground pulse patterns respectively, A, the relative magnitude of the ground pulse was increased from zero to 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5 for a constant value of α, the phase difference between the waves. Readings were taken in this way for values of α at 30-degree intervals from 0 to 360 degrees. Four curves are included in each figure. The uppermost and lowermost curves are for those values of α for which the widest variation of ψ or θ occurred. The other two curves in each figure were included to indicate the rate at which the variation changes for different values of α.

In connection with this study, all of the values for the resultant polarization for $A=0.5$ and $\alpha=0$ to 360 degrees were determined by graphical computation. These values were compared with the instrumentally obtained results. In general, it may be said that the results obtained by use of the instrument are well within the experimental error of the actual ionospheric recording equipment. Values obtained for ψ are within ±6 degrees of the computed value and the values obtained for θ are within ±4 degrees of the computed value.

References

1. Elliptical Polarization and Its Application to the Polarization of Ionospherically Reflected Long Radio Waves, J. M. Kelso. *Technical Report no. 11*, Ionosphere Research Laboratory, State College, Pa., June 1, 1950.

2. The Addition of Elliptically Polarized Waves Including Applications to the Ionospheric Propagation of Low-Frequency Radio Waves, G. H. Friedman. *Technical Report no. 29*, *Ibid.*, Oct. 1, 1951.

3. Waveforms (book), D. Sayre. *Radiation Laboratory Series*, Massachusetts Institute of Technology, Cambridge, Mass., vol. 19, 1949, p. 498.

4. Polarimeter for the Study of Low Frequency Radio Echoes, A. H. Benner, H. J. Nearhoof. *The Review of Scientific Instruments*, New York, N. Y., vol. 21, no. 10, Oct. 1950, pp. 830–34.

5. A Tabulation of 150 Kc/s Polarization Data Obtained at Vertical Incidence Including Their Interpretation from the Single Magneto-Ionic Component Viewpoint, H. J. Nearhoof. *Technical Report no. 25*, Ionosphere Research Laboratory, State College, Pa., Aug. 20, 1951.

◆

Discussion

J. M. Watts (National Bureau of Standards, Washington, D. C.): This paper describes a type of analogue computer. The desired solutions can be quickly arrived at graphically by the traditional circle-diagram method of adding and subtracting 2-dimensional vectors. The accuracy would be as good as, or better than, that obtained by measuring an oscilloscope pattern.

However, it is probable that this instrument would be of some value as a classroom demonstrator in ellipse synthesis. The addition of two voltmeters and a phase meter in parallel with the oscilloscope inputs would improve the accuracy by eliminating the necessity of measurements from the screen. The two methods described for demonstrating the direction of rotation are well known, as is a third which consists of modulating the input voltage with a sawtooth voltage of suitable frequency and amplitude synchronized to a submultiple of the input frequency. Under this system the ellipse changes into an elliptical spiral whose beginning and end show the direction if the sawtooth slope direction is taken into account. This system is particularly easy to apply since the sawtooth sweep oscillator voltage of many oscilloscopes is externally available and may be used independently for the purpose.

G. H. Friedman: I appreciate the comments made by Mr. Watts who is quite correct when he states that more accurate graphical techniques are available for performing computations involving the addition of ellipses. This fact is mentioned in the beginning of the paper. However, in the particular application for which the elliptical polarization synthesizer was designed, close accuracy was neither required nor realistic. The polarization data which formed the basis of the studies being made were scaled from similar elliptical patterns produced on a cathode-ray tube screen by reflected radio waves. Thus, a part of the significance of the instrument described is that it represents an attempt to reduce the burden of trial-and-error computation by presenting the results of that computation in a visual form directly related to experimentally observed phenomena.

Surface and Volume Dielectric Losses

J. J. CHAPMAN
ASSOCIATE MEMBER AIEE

L. F. BLICKLEY
NONMEMBER AIEE

E. A. SZYMKOWIAK
NONMEMBER AIEE

A RELIABLE and convenient system has been devised for measuring large numbers of specimens to detect changes in dielectric constant and dissipation factor due to exposure to controlled conditions of temperature and humidity. A group of four plastic materials consisting of a molded mica-filled phenolic (MFE), a glass silicone laminate (GSG), polytetrafluoroethylene (PF) and polyethylene (PE), were selected for study over the frequency range from 60 cycles per second (cps) to 100 megacycles per second (mc). Measurements of alternating current surface and volume conduction losses, and d-c surface and volume conductivity were made, along with measurements of dielectric strength and flashover strength.[1] The effect of humidity and temperature on these electrical quantities was a major objective in an extensive program. Earlier work at The Johns Hopkins University had greatly stimulated the present interest in the measurement of surface conductivity.[2,3] Although the earlier efforts were made with direct current only, it was generally considered that the work could be extended to the a-c case through the use of a high-voltage Schering bridge. This has been accomplished, data having been obtained at 60 and 1,000 cps. Such measurements are possible with a high-voltage Schering bridge up to 100,000 cps. Using other types of networks, the surface measurements were also taken at 2, 18, and 100 mc. The volume determinations of dielectric constant and dissipation factor were made at the same frequencies. Perhaps the most stringent requirement met in these tasks was that all specimens had to be mounted in holders and measured on location without being disturbed during their 6-month exposures to various conditions of temperature and humidity.

Specimens and Electrodes

To investigate fully both surface and volume phenomena, four types of painted electrode patterns were used, as shown in Fig. 1. Electrode 1, referred to as the

---

Paper 55-81, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, January 31–February 4, 1955. Manuscript submitted October 14, 1954; made available for printing November 30, 1954.

J. J. Chapman, L. F. Blickley, and E. A. Szymkowiak are with The Johns Hopkins University, Baltimore, Md.

This work was performed under Contract *DA-36-039-sc-42478* which was sponsored by the Army Signal Corps. W. F. List and J. S. Smith contributed to the design and development of circuits. Acknowledgment is made to F. L. Schaff and A. Prock for the measurement of the electrical properties of the specimens involved in a lengthy schedule. L. J. Frisco offered timely advice and constructive criticism in accomplishing the work and preparing this report.

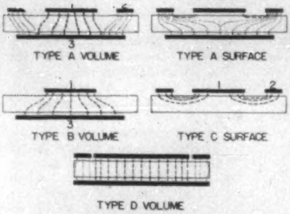


Fig. 1. Electrode patterns on loss specimens showing approximate field shapes

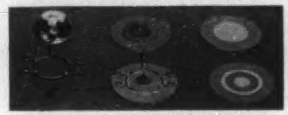


Fig. 2. Type-A loss specimen with stainless steel electrode

measurement electrode, is 1 inch in diameter. Electrode 2, the ring electrode, engages a painted ring which has a $2^1/_4$-inch outside diameter and a 2-inch inside diameter. Electrode 3, the large electrode, is $2^5/_{16}$ inches in diameter.

A type-*A* specimen employs three electrodes which may be connected for either volume or surface measurements. For volume measurements the ring electrode serves as a guard, so its potential is maintained very close in magnitude and phase to that of the measurement electrode. The large electrode is connected to the high-voltage source and the resulting field pattern is somewhat as shown in Fig. 1. The painted patterns and stainless-steel electrodes for a type-*A* specimen are shown in Fig. 2. For surface measurements the potentials of the ring electrode and the large electrode are interchanged, and the large electrode then becomes the guard, while the ring electrode is connected to the high-voltage source. The field pattern shown in Fig. 1 illustrates the approximate stress distribution. The current along the surface may become quite large in a humid atmosphere in the case of some materials. It was found that measurements made with direct current, 60 cps, 1,000 cps, and 2 mc gave reliable results for both surface and volume with the type-*A* specimens. At higher frequencies it was necessary to use two electrode specimens because capacitance effects made it impractical to maintain a third electrode at the proper potential.

The type-*B* specimen employs electrodes 1 and 3, the latter being connected to the high-voltage source. This type of specimen is suitable for the measurement of volume properties from 2 mc to 100 mc. A guard ring is not required for the high-frequency volume measurements because the effects of surface loss have proved negligible for such an arrangement. At high frequency even a leakage path of 10 megohms would have a negligible effect when it is associated with a dielectric part which has a capacitive reactance of perhaps 100 ohms. The Q of such a component would be 10^5 and therefore would not be likely to compromise the functioning of a circuit. By contrast, a 10-megohm leak across the same capacitive element at 60 cps would cause a very large phase defect angle in the circuit. In the case of materials which had no measurable surface conduction, the type-*B* sample could be used to check the volume measurements of the type-*A* sample at 60 cps, 1,000 cps, and 2 mc.

The type-*C* sample employs electrodes 1 and 2 and is suitable for surface measurements at 18 and 100 mc. In this case the ring electrode is connected to the high-voltage source. This type of specimen also provided a qualitative check on the surface measurements of the type-*A* specimen at 60 and 1,000 cps.

The use of these three types of specimens made it possible to cross check various observations pertaining to the surface and volume behavior of the four materials. Because the field patterns involved in the *A*, *B* and *C* specimens are admittedly complicated, a considerable amount of calibration was necessary. For this purpose the type-*D* specimen was used to evaluate all materials at 60 and 1,000 cps. This conventional specimen, shown in Fig. 1, has a 3-inch diameter measurement electrode with a very close fitting guard ring, thereby affording the simplest geometrical configuration. In this manner all types of specimens were reduced to an equivalent geometry at 60 cps, which was confirmed at 1,000 cps. This geometric equivalence is independent of frequency over the chosen frequency range, since the specimen dimensions were insignificant compared to wave length, even at 100 mc.

Conditioning Cabinets

Storage under nine conditions of exposure took place in specially constructed cabinets the lids of which were equipped to hold 36 specimens for loss measurements, so that variations in properties could be followed without interruption for periods up to 6 months. Fig. 3 shows a typical lid with various types of specimens mounted; actually six positions were usually taken up with air capacitors for the purpose of checking measuring equipment. All such lids were equipped with rectangular screen baskets, to shield the specimens from stray electrical effects during measurement. This entire assembly was placed on a rubber gasket which provided guard insulation and sealed the cabinet as well.

Fig. 3. Conditioning cabinet lid with loss specimens mounted

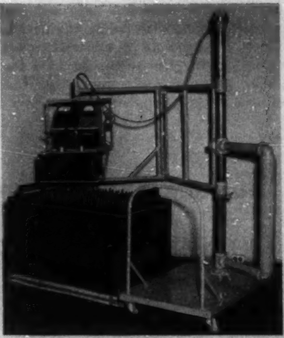


Fig. 4. Conditioning cabinet and mobile measurement unit

The cabinets had an additional 4 cubic feet of storage space below the lid and guard screen for other types of speci-

Table I. Nine Conditions of Exposure

Temperature, C	Per-Cent Relative Humidity (RH)
−55	unspecified
25	20
25	50
25	100
50	20
50	50
50	100
85	unspecified
125	unspecified

Table II. Dielectric Constant K and Dissipation Factor D Initial Values

Temperature, C	60 Cps		1 Kc		2 Mc		18 Mc		100 Mc	
	K	D	K	D	K	D	K	D	K	D
PE										
−55	2.35	L	2.35	L	2.31	L	2.30	L	2.26	L*
25	2.31	L	2.31	L	2.22	L	2.22	L	2.19	L
50	2.27	L	2.27	L	2.19	L	2.19	L	2.19	L
PF										
−55	2.15	L	2.15	L	2.14	L	2.12	L	2.08	L
25	2.10	L	2.10	L	2.10	L	2.10	L	2.08	L
50	2.09	L	2.09	L	2.08	L	2.08	L	2.05	L
85	2.02	L	2.02	L	2.00	L	2.00	L	2.00	L
125	2.02	L	2.02	L	2.01	L	2.00	L	2.00	L
MFE										
−55	4.45	0.018	4.45	0.010	4.59	0.004	4.58	0.003	4.25	0.018
25	5.90	0.125	5.40	0.120	4.80	0.015	4.70	0.012	4.45	0.012
50	7.50	0.275	5.95	0.120	4.90	0.015	4.82	0.012	4.63	0.024
85	14.70	0.411	8.80	0.405	4.82	0.015	4.80	0.010	4.65	0.041
125	15.10	0.480	15.10	0.450	5.20	0.015	5.10	0.010	4.75	0.045
GSG										
−55	3.89	L	3.89	L	3.78	L	3.78	L	3.73	0.010
25	4.00	0.003	3.97	0.004	3.90	0.004	3.88	0.002	3.87	0.006
50	3.90	0.003	3.87	0.004	3.85	0.004	3.85	0.002	3.85	0.006
85	3.67	0.005	3.60	0.005	3.60	0.002	3.60	0.002	3.60	0.005
125	3.65	L	3.65	L	3.57	0.002	3.57	0.002	3.57	0.005

*L = less than 0.001.

mens. Fig. 4 shows a cabinet under the mobile measurement equipment. The nine conditions of exposure are shown in Table I.

All cabinets were copper-lined and equipped with circulating blowers and temperature controls. Humidity in some cases was controlled by maintaining a layer of water in the bottom of the box at the desired wet-bulb temperature. For the 25C, 20-RH condition, it was necessary to tap a dry compressed air supply and feed one of the cabinets at a slow but steady rate. Conditioning at −55C was accomplished by use of a 2-stage-compressor deep-freeze cabinet. The lid containing the loss specimens was constructed in a manner similar to all others except that it was 2 inches thick and was extremely well insulated thermally. Except during periods of measurement, the outer metal-clad lid of the unit was kept closed. Appreciable frosting at the electrical entry points caused no serious difficulty during measurement procedure, since frost apparently has a reasonably high dielectric strength.

Methods

The high-voltage Schering bridge used during the investigation was constructed for the particular task. Although regular routine called only for measurements at 60 and 1,000 cps, particular points of interest sometimes caused a specimen to be traced up to 100,000 cps with this device. It was found that the flexible guarded cables to the specimens could be 8 feet long without sign of error. The bridge is conventional in most electrical details except that the guard circuit was of unusually low impedance both in its balancing network and high-voltage coupling. This circuit was made "stiff" to offset any effects of surface leakage on guard insulation in the more humid conditioning cabinets. Since the specimens had to suit higher frequency purposes as well, their capacitance was of the order of 10 micromicrofarads. Some accuracy was sacrificed by using such small specimens; with 1,000 volts applied to the bridge, the lower limit of dissipation factor detection was 0.001.

At 2, 18, and 100 mc another type of network was used for the measurement of dielectric constant and dielectric loss of the specimens in the cabinets. This network for some whimsical reason became known as the Beechnut circuit; it has been described in its preliminary development.[4] Inherently it has about the same sensitivity as a *Q*-meter, but rather than measure resonant rise of voltage in a series circuit, it measures residual current in a parallel resonant circuit. Its great advantage is that in working around conditioning cabinets, all metal enclosures surrounding the specimens become an approximate low-voltage guard and thus intercept loss currents arising from wet or faulty bushings. In extending the use of the device up to 100 mc, it was found advisable to make the adjustable quantity in the parallel resonant circuit the inductance rather than the capacitance.

An important feature in the use of this circuit at 2, 18, and 100 mc, was mobility. The network had to be brought close to the specimens mounted in each conditioning cabinet. Fig. 4 shows a 2-mc Beechnut network suspended and in position for a series of measurements. The mechanics of connection to each specimen terminal are quite exacting. Contact resistance must be avoided at these terminals and also at the junction surfaces of the mounted specimens. Errors leading to exceptionally high values of dissipation factor at 100 mc were ultimately traced to contact resistance. The air capacitors in each conditioning cabinet covered the range of all measurement devices so that test location or grounding condition could not change capacitive calibration without notice. Further, these capacitors provided a zero loss reference at all frequencies.

The voltage applied to the 1/8-inch thick specimens, in measuring dielectric constant and dissipation factor was 1,000 volts at 60 cps but less at higher frequencies. At 2 and 18 mc it was 125 volts, while at 100 mc it was kept at 50 volts. This tapering of voltage with respect to frequency was consistent with maintaining ample sensitivity and at the same time avoiding changing the internal temperature of a specimen during measurement. At 2, 18, and 100 mc the power supplies were about 60-watt output. At lower frequencies the bridge supply was capable of delivering over 1,000 watts.

Volume Properties

Table II gives the initial values of dielectric constant and dissipation factor of the four materials over a wide range of temperature and frequency. In many cases 1 hour elapsed in letting the specimens come to temperature equilibrium. Using the rather small specimens previously described, PE and PF showed no detectable dissipation factor initially or with prolonged exposure to various conditions. Among laminated materials, GSG may be considered a low loss material throughout the range of initial tests.

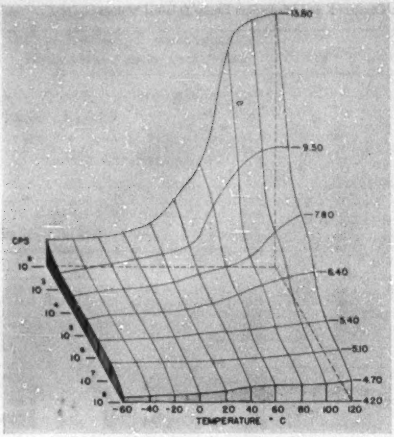


Fig. 5. Effect of temperature on the dielectric constant of MFE

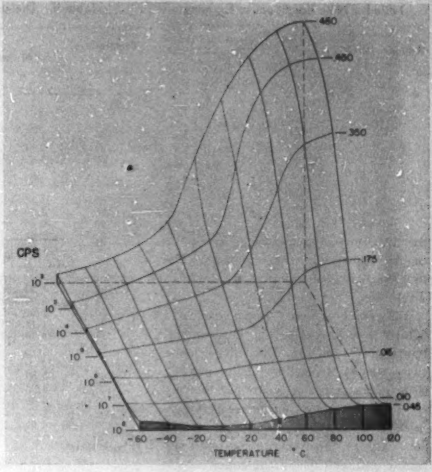


Fig. 6. Effect of temperature on the dissipation factor of MFE

It is probable that this material lost a small amount of residual moisture during these tests to lower further the values of dissipation factor. The values of Table II provide interesting 3-dimensional plots of dielectric constant and dissipation factor for MFE in Figs. 5 and 6 respectively. The enormous changes with temperature exhibited by MFE are moderated if the material is subjected to a preliminary drying treatment. The properties given in this table pertain to materials that were stored for several months at 25C, 45 RH, after delivery from the manufacturer. MFE loses a significant amount of moisture when stored in the presence of a dessicant; further, it gains considerable weight in a humid atmosphere.[1]

Prolonged exposure to high humidity produced drastic changes in the electrical properties of MFE and GSG. The 3-dimensional plots of Figs. 7 and 8 illustrate the effect of exposure at 25C, 100 RH on MFE; deleterious trends of similar magnitude are also obtained at 50C, 100 RH. Figs. 9 and 10 illustrate the effects of exposure at 50C, 100 RH, on the dielectric constant and dissipation factor of GSG. Effects of similar magnitude are obtained at 25C, 100 RH.

The presence of moisture in these di-

Vol. 7
Pt. 1
1955

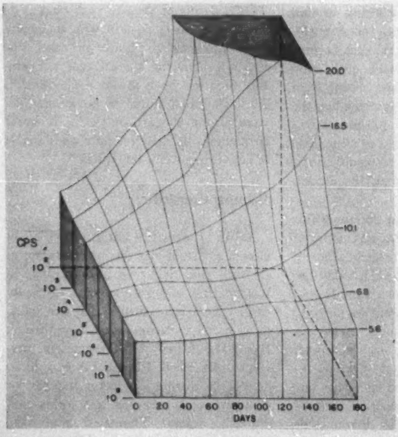


Fig. 7. Effect of aging at 25C, 100 RH, on the dielectric constant of MFE

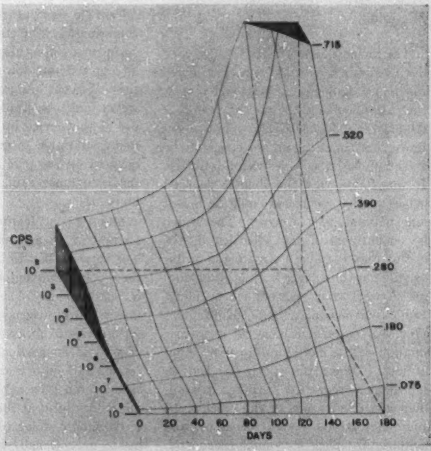


Fig. 8. Effect of aging at 25C, 100 RH, on the dissipation factor of MFE

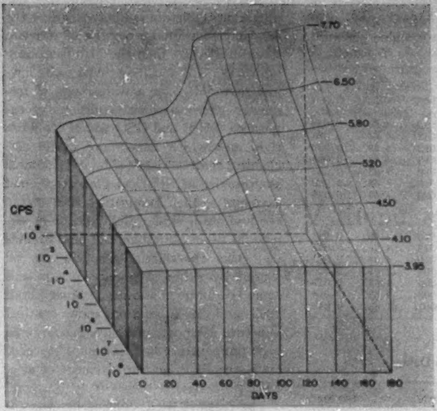


Fig. 9. Effect of aging at 50C, 100 RH, on the dielectric constant of GSG

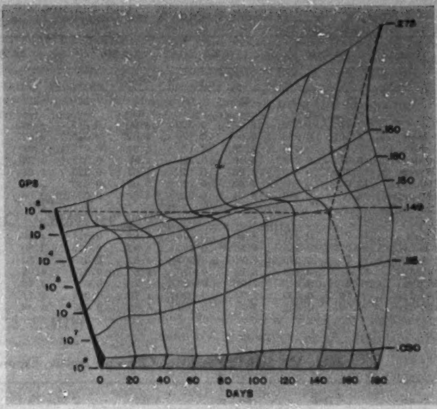


Fig. 10. Effect of aging at 50C, 100 RH, on the dissipation factor of GSG

electrics is most noticeable at low frequency. This appears to be the case whether the moisture is residual from the manufacturing process or admitted through exposure to a humid atmosphere. The detection of moisture by means of measuring resistivity with a d-c source and galvanometer, proved to be simpler and even more sensitive than the low-frequency measurements. This is illustrated by comparing the d-c volume resistivities of MFE in Table III, and the corresponding a-c properties in Table IV. The a-c measurements were consistent with the trends of Table III to the extent of showing improved electrical properties under the drying conditions, and extensive deterioration at high humidity. Although the determination of d-c resistivity provides a simple but approximate means of evaluating materials quite generally as to moisture absorption, the exact determination of a-c properties is often necessary. Part of the objective in this investigation of the electrical properties of materials has been to relate dielectric breakdown with variations in dielectric constant and dissipation factor; the complementary relation of the 3-dimensional diagrams of these properties at various conditions is quite definite.[1]

At room temperature PE, PF, and GSG were found to have resistivities greater than 10^{16} ohm-cm, while thoroughly dried MFE had a value of 5×10^{14} ohm-cm. These high, but approximate, values of resistivity were obtained measuring the decay time constants of charged specimens. The limit of detection using a Leeds and Northrup insulation resistance test set and the specimens previously described was 10^{14} ohm-cm volume resistivity.

The effects of exposure at 25C, 20 RH, and 25C, 50 RH produced no noteworthy changes in the four materials examined over a period of 6 months. Likewise, the condition −55C proved to have no significant effect, other than the initial change in properties due to temperature change.

Surface Properties

In the study of surfaces, no definite leakage could be reported for PE or PF at any of the conditions. This was true at all frequencies and with direct current. Prior to this experimental work, it had been part of the speculation, that even in the case of surfaces that did not wet, individual water droplets would be coupled into high-frequency fields, and thus exhibit ionic and interfacial loss. The results on very good dielectric materials, such as PE and PF, were inspected for such effects, but no confirming surface losses were observed. In quite dry air, surface leakage for MFE and GSG was too small to detect at all temperatures. The d-c equipment was most sensitive in detecting surface leakage, reaching its limit at 10^{14} ohms per square. In the case of the materials examined it is quite probable that in a dry atmosphere the surface values were perhaps two decades beyond this range. At room condition, or at higher humidities, GSG and MFE exhibited measurable amounts of surface leakage. As indicated by Table V, the range of detection was dependent upon frequency; at 60 cps, the limit was 10^{13} ohms per square and this value decreased with increase in frequency. In

Table III. D-C Volume Resistivity, Ohm-Cm, Exposure Data

Time	25C, 100 RH	50C, 100 RH	50C, 50 RH	85C	125C
			MFE		
0	1.2×10^{13}	1.3×10^{13}	1.3×10^{13}	5.3×10^{12}	5.3×10^{13}
1 hour	7.8×10^{12}	3.6×10^{11}	1.1×10^{13}	5.3×10^{13}	5.3×10^{13}
1 day	2.4×10^{12}	1.3×10^{11}	5.5×10^{12}	5.3×10^{13}	5.3×10^{13}
10 days	1.2×10^{11}	5.5×10^{10}	2.6×10^{11}	6.5×10^{13}	1.0×10^{14}
50 days	3.4×10^{9}	6.3×10^{8}	1.4×10^{11}	1.9×10^{13}	2.5×10^{13}
100 days	4.5×10^{8}	3.7×10^{7}	9.9×10^{10}	G*	G
180 days	1.1×10^{7}	1.5×10^{7}	1.8×10^{10}	G	G
			GSG		
0	G	G	G	G	G
1 hour	8.6×10^{10}	7.3×10^{10}	8.0×10^{13}	G	G
1 day	4.7×10^{10}	3.4×10^{10}	8.0×10^{13}	G	G
10 days	1.4×10^{10}	1.2×10^{10}	7.9×10^{13}	G	G
50 days	4.2×10^{9}	2.6×10^{9}	3.5×10^{13}	G	G
100 days	2.0×10^{9}	1.2×10^{9}	4.6×10^{13}	G	G
180 days	1.2×10^{9}	2.8×10^{8}	4.6×10^{13}	G	G

*G = greater than 10^{14}.

Table IV. MFE Dielectric Constant K and Dissipation Factor D, Exposure Data

Time	60 Cps		1 Kc		2 Mc		18 Mc		100 Mc	
50C, 50 RH										
	K	D	K	D	K	D	K	D	K	D
0	7.50	0.275	5.95	0.120	4.90	0.015	4.82	0.012	4.63	0.024
1 hour	7.50	0.275	6.06	0.120	4.92	0.015	4.82	0.012	4.63	0.024
1 day	7.50	0.275	6.28	0.120	4.94	0.015	4.87	0.012	4.63	0.024
10 days	9.07	0.290	6.49	0.120	4.96	0.022	4.90	0.015	4.63	0.024
50 days	11.30	0.310	7.03	0.140	5.03	0.025	5.03	0.020	4.63	0.024
100 days	12.00	0.330	7.03	0.150	5.03	0.030	5.03	0.020	4.63	0.030
180 days	15.5	0.330	8.61	0.220	5.06	0.030	5.06	0.020	4.63	0.035
85C										
0	14.7	0.411	8.80	0.405	4.82	0.015	4.80	0.010	4.65	0.041
1 hour	14.7	0.250	8.80	0.405	4.82	0.015	4.80	0.010	4.65	0.041
1 day	14.7	0.250	8.80	0.060	4.82	0.015	4.80	0.010	4.65	0.041
10 days	12.8	0.180	8.34	0.030	4.80	0.013	4.78	0.009	4.65	0.030
50 days	10.2	0.035	7.88	0.025	4.75	0.011	4.75	0.009	4.63	0.015
100 days	9.75	0.030	7.75	0.020	4.73	0.009	4.71	0.008	4.63	0.015
180 days	9.69	0.025	7.71	0.010	4.68	0.008	4.68	0.007	4.63	0.015

Table V. Surface Resistivity, Ohms per Square, Exposure Data

Time	D-C	60 Cps	1 Kc	2 Mc	18 Mc	100 Mc
MFE, 25C, 100 RH						
0	1.7×10^{13}	G_1*	G_1	G_2†	G_3‡	G_4§
1 hour	9.5×10^{12}	7.2×10^{11}	G_1	G_2	G_3	G_4
1 day	7.1×10^{12}	6.0×10^{11}	1.0×10^{11}	5.0×10^{7}	5.0×10^{6}	G_4
10 days	8.4×10^{10}	5.2×10^{9}	3.3×10^{8}	3.9×10^{7}	5.0×10^{6}	G_4
50 days	3.2×10^{8}	2.9×10^{7}	8.5×10^{7}	1.0×10^{6}	6.3×10^{5}	3.0×10^{5}
100 days	1.4×10^{8}	7.0×10^{6}	2.4×10^{7}	6.0×10^{5}	1.6×10^{5}	8.0×10^{4}
180 days	1.1×10^{7}			5.2×10^{5}	8.5×10^{4}	5.8×10^{4}
GSG, 25C, 100 RH						
0	2.5×10^{12}	2.2×10^{11}	1.0×10^{11}	G_2	G_3	G_4
1 hour	5.5×10^{10}	2.2×10^{10}	2.5×10^{10}	2.5×10^{7}	6.0×10^{6}	G_4
1 day	1.1×10^{10}	4.0×10^{9}	3.4×10^{9}	1.5×10^{7}	6.0×10^{6}	G_4
10 days	2.6×10^{9}	1.5×10^{9}	1.8×10^{7}	1.0×10^{6}	8.0×10^{5}	G_4
50 days	2.6×10^{8}	4.0×10^{8}	1.5×10^{7}	7.5×10^{5}	6.0×10^{5}	2.0×10^{5}
100 days	2.6×10^{8}	1.5×10^{8}	8.5×10^{6}	5.5×10^{5}	1.4×10^{5}	2.0×10^{5}
180 days	2.5×10^{8}	7.5×10^{6}	3.0×10^{6}	4.0×10^{5}	1.4×10^{5}	1.9×10^{5}
GSG, 50C, 50 RH						
0	1.7×10^{12}	2.2×10^{11}	G_1	G_2	G_3	G_4
1 hour	1.7×10^{12}	2.2×10^{11}	G_1	G_2	G_3	G_4
1 day	2.0×10^{12}	4.0×10^{10}	G_1	G_2	G_3	G_4
10 days	3.9×10^{12}	6.4×10^{9}	G_1	G_2	G_3	G_4
50 days	4.0×10^{12}	1.4×10^{9}	G_1	G_2	G_3	G_4
100 days	5.5×10^{11}	1.4×10^{9}	G_1	G_2	G_3	G_4
180 days	9.5×10^{12}	1.4×10^{9}	G_1	G_2	G_3	G_4

*G_1 = greater than 10^{13}.
†G_2 = greater than 10^{8}.
‡G_3 = greater than 10^{7}.
§G_4 = greater than 5×10^{5}.

Table V, it is possible to trace a decrease in surface resistivity across the frequency range for either MFE or GSG at a particular exposure time at 25C, 100 RH. In some instances this trace is interrupted by lack of sensitivity at the high-frequency end.

Both these materials show surface leakage at 50C, 100 RH; their behavior in this case is not greatly different from that at 25C, 100 RH. Like many other glass laminates and ceramics, GSG is most susceptible to surface leakage. MFE does not wet to as great an extent and shows no effect at 50C, 50 RH. Table V indicates that GSG does show some effects at this condition, although they are of a sporadic nature and limited to d-c and 60 cps observations. This property was examined at all conditions, but without noteworthy result except for the foregoing condition involving high densities of water vapor.

Table V indicates that at 25C, 100 RH prolonged exposure continues to cause a decrease in surface resistance in the case of GSG and MFE. Such a conclusion may be drawn, but cognizance of the ultimate increase in volume losses must be considered as possibly modifying the observation to some extent. A thin layer of insulation near and adjoining the surface may play a part in the surface measurement.

Conclusions

Increased loss attributable to moisture is more readily detected at 60 cps and direct current than at radio frequency. This is true in the case of both volume and surface measurements. Accompanying the low frequency region of sensitive detection is the more critical problem of application; moisture is likely to affect the low-frequency or d-c circuits of practice more drastically then the high-frequency networks. PE and PF proved to be low-loss materials over a wide range of frequency under various conditions of exposure. MFE and GSG which are more rigid and which have more useful mechanical characteristics show considerable compromise of electrical properties upon exposure to high humidity. The surface conduction of GSG quickly becomes significant in an atmosphere of high humidity; this characteristic appears to be shared by other glass fabrications and some ceramic materials. Since both GSG and MFE contain some residual moisture after manufacture, exposure to high temperature but low humidity generally results in an improvement of electrical properties.

The method used, considering the coordinated construction between storage cabinets and measurement devices as a whole, was most satisfactory for the large program. The feature of including six air capacitors in each conditioning cabinet covering the range of probable specimen values was most useful. Such capacitors are most crude in construction, being very approximately spaced, but serve adequately as standards after being evaluated by the high-voltage Schering bridge at 60 cps.

References

1. Dielectric Failure of Volume and Surface Types, **J. J. Chapman, L. J. Frisco, J. S. Smith.** *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 349–56.

2. Measuring of Surface and Volume Resistances, **R. K. Witt, J. J. Chapman, B. L. Raskin.** *Modern Plastics*, New York, N. Y., April 1947, pp. 151–53.

3. A Method for Evaluating the Surface and Volume Resistance Characteristics of Solid Dielectric Materials, **W. G. Amey, F. Hamburger, Jr.** *Proceedings*, American Society for Testing Materials, Philadelphia, Pa., vol. 49, 1949, pp. 1079–91.

4. A Guarded Circuit for Measurement of Dielectric Properties at Radio Frequencies, **J. W. Dzimianski, R. K. Witt, J. J. Chapman.** *Annual Report*, Conference on Electrical Insulation, National Research Council, Washington, D. C., 1948, pp. 8–10.

◆

Discussion

G. M. L. Sommermann (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The conclusion that moisture affects the electrical properties more at d-c and 60-cps than at radio frequencies confirms results obtained in other investigations on insulating paper and on nylon films. The first increments of water absorbed by the material exposed to moist air are held as tightly bound absorbed molecules (if not in chemical combination); these may not greatly increase even the low-frequency dielectric constant, as shown in Fig. 9. Additional increments of water absorbed during further exposure are held as less tightly bound molecules which do increase the dielectric constant substantially at low frequencies but, in Fig. 9, are not free enough to contribute to the dielectric constant at 100 mc.

J. J. Chapman, L. F. Blickley, and E. A. Szymkowiak: Mr. Sommerman points to a conclusion that is of considerable practical interest. If the problem were to detect absorbed moisture in solid insulation perhaps for plant control purposes, the d-c or 60-cyps type of measurement provides a simple and economical means. The problem of moisture in insulating compounds is so universal that detection and control are necessary in dealing with solids and impregnants. Two factors are to be considered in control work, sensitivity and reproducibility. For a particular line of materials or apparatus, the proper compromise between these two factors may be found with d-c or at some low a-c frequency.

1, 74
1
55

Dielectric Failure of Volume and Surface Types

J. J. CHAPMAN
ASSOCIATE MEMBER AIEE

L. J. FRISCO
ASSOCIATE MEMBER AIEE

J. S. SMITH
ASSOCIATE MEMBER AIEE

IN THIS investigation, an attempt is made to obtain an interpretation of phenomena and also to serve practical ends. Since laboratory tests are preferably simple and of a controlled nature, the results are very limited when weighed against the multitudinous considerations of field applications. After earlier exploration, in which both materials and techniques were given careful scrutiny, four solid dielectrics were chosen for the present study.[1–3] Failure of practical dielectric structures in most cases is clearly through a volume of solid or else across the surface at an interface. Therefore, values of electric stress necessary to produce puncture constitute many of the compiled data in this paper. Likewise, flashover voltage is given for the materials as determined from the use of a fixed structure in air. The frequency range was from 60 cycles per second (cps) to 100 megacycles per second (mc), and the effects of temperature and humidity were studied.

Specimens

The four materials studied were molded mica-filled phenolic plastic (MFE), laminated glass silicone (GSG), polytetrafluorethylene (PF), and polyethylene (PE) sheet stocks. Six values were recorded for every flashover or breakdown determination in covering all of the conditions of the program. The particular type of high-voltage electrode used in the dielectric strength tests was a brass tube, 7/16 inches in outer diameter and 5/16 inches in inner diameter; the inner and outer edges were rounded to a radius of 1/64 inch. The electrode was inserted into a cavity of 7/16-inch diameter and approximately 3/8-inch depth, as shown in Fig. 1. This deep cavity, of course, required the use of 1/2-inch thick specimens; the bottom of the cavity contained a groove to fit the tube and of such additional depth as to allow only a small breakdown distance along the line *CD* to the other side of the specimen. The breakdown distance for MFE, PF, and PE was 30 mils, while GSG was tested at 60 mils. At these small distances it was found that this type of electrode, and other types as well, gave values of dielectric strength corresponding to a uniform field as long as all edges were rounded to at least 1/64-inch radius. The cavities were sprayed with a conducting silver paint, which was engaged by the inserted electrode. More recently in checking by other methods, it has been found that a 1/2-inch diameter twist drill forms a very smooth cavity and that the bottom is tapered suitably for breakdown tests. Moreover, sprayed solder forms an ideal conducting coating for such cavities.[4] The use of the standard drill is most convenient, except in the case of very hard materials where tubular electrodes may be ground into the desired depth without difficulty. Recessed high voltage electrodes are usually necessary in dielectric strength tests to avoid arcing in in the immersion liquid of the breakdown cell. The bottom face of the specimen must also be coated with a conducting film which engages a large flat electrode at ground potential.

The flashover voltage for each material was obtained across a 3/8-inch gap between stainless steel washers. Fig. 2 illustrates the method of assembly, where a piece of 1/8-inch stock, $2^{3}/_{4}$ inches long and $1^{1}/_{4}$ inch wide is used. Stainless steel hardware was chosen to avoid metallic corrosion during long periods of exposure at high humidity. Fig. 3 shows a specimen mounted for test at controlled humidity and temperature.

Conditioning and Testing of Specimens

Sealed cabinets for storage of specimens for periods as long as 6 months were a necessary part of the equipment. These were all located in an area maintained at 25 degrees centigrade (C) and approximately 45-per-cent relative humidity (RH). Control of the ambient condition thus made the maintenance of specified conditions within the various cabinets an easier task. Additional features of these storage facilities are discussed in another paper.[5] Nine conditions were studied, as shown in Table I. Since PE softens in the vicinity of 85C, storage conditions 8 and 9 did not include such specimens.

While various jacketed test cells, both for dielectric strength and flashover tests, were necessary for the departures from ambient conditions, their tedious construction and particular usage merit little description. In general, it was found advantageous to resort to heat exchanger schemes at higher frequencies, so as to

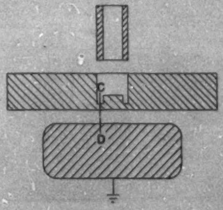


Fig. 1. Cross-sectional view of puncture sample with recessed, tubular electrode

---

Paper 55-82, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31-February 4, 1955. Manuscript submitted October 14, 1954; made available for printing November 30, 1954.

J. J. Chapman, L. J. Frisco, and J. S. Smith are with The Johns Hopkins University, Baltimore, Md.

This work was performed under contract *DA-36-039-sc-42478*, which was sponsored by the Army Signal Corps. The authors are indebted to their co-workers, W. G. Baumann, R. D. Picard, and E. A. Szymkowiak for their part in conducting the high-voltage tests.

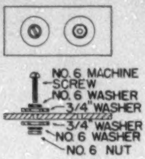


Fig. 2. Flashover sample assembly

Fig. 3. Conditioning chamber for flashover tests

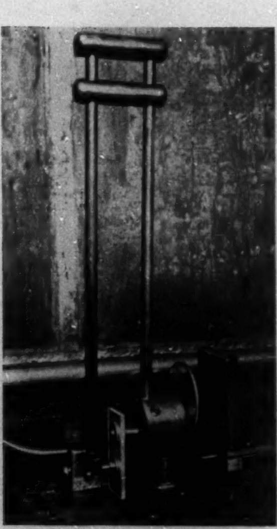


Fig. 4. Resonant circuit for 100-mc flashover tests

keep all heaters, motors, and control circuits at a minimum in the immediate vicinity of high-power resonant circuits. For short-time cooling to −55C at test locations liquid nitrogen was ideal, since the flow through a heat exchanger could be controlled at remote location. At 25C and over, transformer oil was used as the immersion medium in breakdown cells; at −55C it was necessary to use Varsol, which remains in the liquid state. For flashover tests where humidity as well as temperature was a factor at test location, air was conditioned and then circulated by a blower through ducts around the specimens. At 100-mc test location, the piping and duct work was more extensive because all control elements and accessories had to remain outside of a shielded room.

In the case of both dielectric breakdown tests and flashover tests, the voltage was applied gradually and failure was accomplished in approximately 40 seconds. Apparatus suitable for developing and measuring breakdown voltages has been described in previous publications.[1-3] It is worthy of mention here, that at 100 mc the breakdown cell had to be fashioned compactly to suit the specimen.[2] The resonant circuit for the puncture and flashover tests had to be made specially for each type of test specimen. Fig. 4 shows the 100-mc resonant circuit used in flashover tests; the vertical rods with fixed and movable slide bars form a coarsely adjustable inductance. The input power is fed into an increment of the left rod near the ground plane. The terminal connected to the specimen is quite large and forms a capacitance junction for the fine tune extension of the ground plane at *A* and the "line-of-sight" probe directly below.

Results of Dielectric Strength Tests

Figs. 5 through 8 are 3-dimensional diagrams which summarize the effects of temperature and frequency on the dielectric strength of the four materials. These diagrams represent the initial evaluation of materials received from manufacturers, although at least 1 hour was permitted at extreme temperature (high or low) for purposes of equilibrium. The temperatures shown represent those of the materials immediately prior to the application of voltage. PE was held at 80C for a necessarily short period because this temperature is near its softening point.

In the case of MFE, the character of the failure changes with the increasing frequency of the applied voltage. At low frequencies, the puncture path is very narrow and well defined but at radio frequencies the failure includes a broad volume that swells and is thoroughly weakened mechanically. Although considerable heat is evolved in the high-frequency type of failure, no charring is evident. GSG shows a rather narrow puncture path up to 100 mc. At this high-frequency, delamination sometimes accompanies the formation of the path. PE and PF exhibit the narrow, definite failure paths at all frequencies.

Table I. Nine Storage Conditions

No.	Temperature, C	RH
1	−55	unspecified
2	25	20
3	25	50
4	25	100
5	50	20
6	50	50
7	50	100
8	85	unspecified
9	125	unspecified

Although all 9 conditions of exposure were pursued for 6 months in the case of all materials, no significant change in dielectric strength can be reported in the case of PE and PF. The only significant decreases in dielectric strength occurred in the case of MFE and GSG, and these suffered deterioration only at high humidity. Prolonged exposure to high or low temperature showed no adverse electrical effects. In fact, prolonged exposure at 85 and 125C apparently drove some residual moisture from MFE and GSG and improved the dielectric strength slightly.

Figs. 9 and 10 demonstrate the adverse effect of moisture absorption on the dielectric strength of MFE and GSG respectively, and indicate that the effect of moisture shows markedly at 60 cps. As MFE and GSG absorb moisture the character of the failure changes. The broad, volume-type of failure occurs at progressively lower frequencies as the period of exposure to high humidity is increased, and finally even at 60 cps the end point is not sharply defined.

Complete plots of the 50C, 100-RH exposure data have been made and they differ only slightly from the 25C, 100-RH diagrams of Figs. 9 and 10. The final values of dielectric strength after 180 days of exposure to 50C, 100 RH are lower than the corresponding values at 25C, 100 RH. Table II shows the change in dielectric strength of MFE and GSG after 180 days of exposure at 50C, 50 RH and at 50C, 100 RH. The final values are compared to the initial 50C values at each frequency. Although the 50C, 50-RH condition caused a significant change in the dielectric strength of both materials,

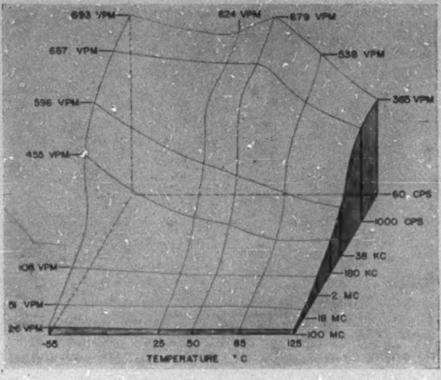


Fig. 5. Effect of temperature on dielectric strength of MFE

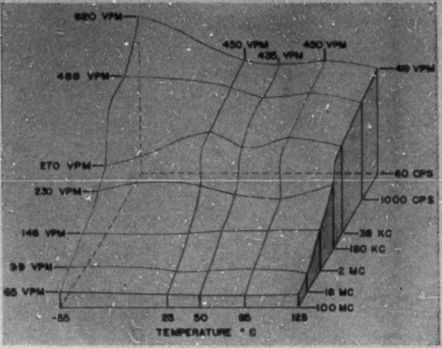


Fig. 6. Effect of temperature on dielectric strength of GSG

the effect is much less drastic than that of the 100-RH condition.

Concurrent Changes in Dielectric Constant and Dissipation Factor

The complete discussion of the associated measurements of dielectric constant and dissipation factor of MFE, GSG, PF and PE is the subject of another paper.[5] At this time, however, some results on GSG, for example, appear to be relevant to a discussion of the dielectric breakdown results. Fig. 11, which is a complete plot of dissipation factor in the exposure study at 25C, 100 RH, is in many ways complementary to the corresponding plot of Fig. 10.

The plane indicating the initial variation of dissipation factor with frequency, Fig. 11, shows the value decreasing to a minimum between 10^6 and 10^7 cps. This value is about 0.002, but as the frequency is increased to 100 mc, the value has increased to 0.006. The component of loss which increases with frequency conceivably is present on a minute scale at 10^5 cps, but is obscured by the decreasing component. Such deviations in the polarization mechanism may be pictured as distributed in the frequency spectrum and affecting the drastic decrease in breakdown strength, first evident at approximately 10^5 cps (point *A*, Fig. 10). Greater credence is given this idea by further inspection of Figs. 10 and 11. As moisture is absorbed as a result of exposure, the decreasing and increasing components of dissipation factor become less distinguishable as functions of frequency. It may be presumed that such components now have broad distribution functions and that the region of overlap extends to low frequency. The shape of

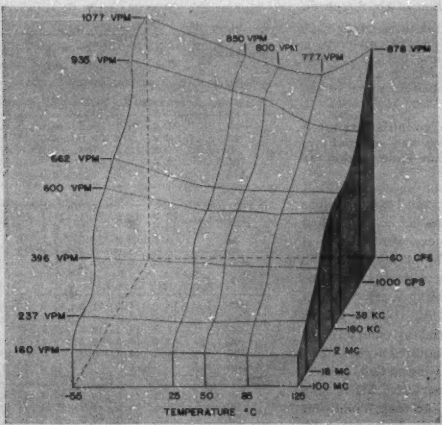


Fig. 7. Effect of temperature on dielectric strength of PF

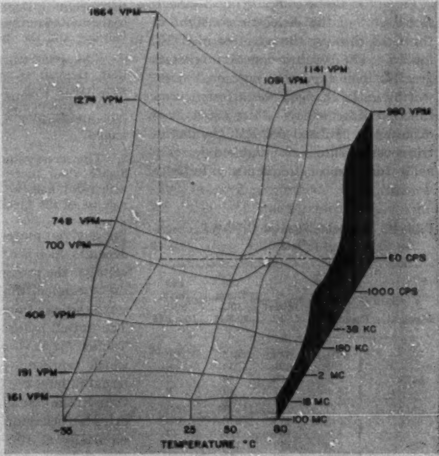
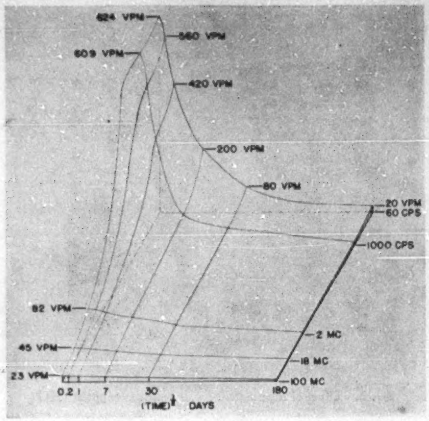


Fig. 8. Effect of temperature on dielectric strength of PE



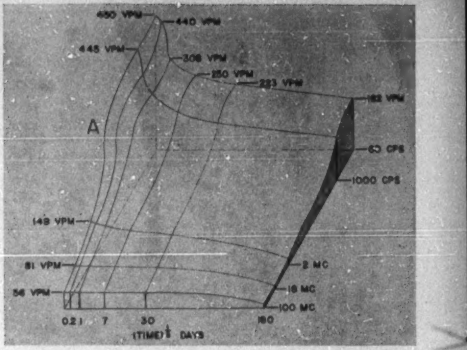


Fig. 9 (left). Effect of aging at 25C, 100 RH, on dielectric strength of MFE

Fig. 10 (above). Effect of aging at 25C, 100 RH, on dielectric strength of GSG

the dielectric strength curve as a function of frequency has changed drastically, a marked decrease occurring in the region of 10^3 cps after 180 days. What is proposed is that, although the detection of a loss mechanism, which increases with the applied frequency, predicts thermal phenomena at higher frequency, the exceptional dynamics of electrons in the locality of deviation moments cause a decrease in breakdown strength. The greatest decrease in dielectric strength as a function of frequency occurs before the dissipation characteristic is noticeably increasing with respect to frequency.

Fig. 12 is the accompanying 3-dimensional plot of the dielectric constant of the GSG showing the effect of high humidity. The dielectric constant increases greatly upon exposure, but most noteworthy is the changed characteristic as a function of frequency. After 180 days of exposure at 25C and 100 RH, the dielectric constant shows very marked decrease as a function of frequency. Initially there was no significant difference in value of dielectric constant between 60 cps and 10^5 cps while, after 180 days, the difference is evident between 60 cps and 1,000 cps.

Table II. Dielectric Strength, VPM, Exposure Data

Material	Frequency	Initial 50C	180 Days, 50C, 50 RH	180 Days, 50C, 100 RH
MFE	60 cps	679	615	10
	1 kc	609	530	6
	2 mc	78	50	3
	18 mc	46	29	2
	100 mc	20	18	1
GSG	60 cps	435	310	156
	1 kc	404	273	42
	2 mc	159	134	26
	18 mc	87	81	21
	100 mc	53	46	10

Results of Flashover Tests

Using the electrode configuration shown in Fig. 2, the flashover test can result in the following types of failures:

1. The stress within the material is well below the breakdown level and the failure occurs in the air gap between the washers. The resulting arc may permanently damage the specimen, depending on the properties of the material, the frequency of the applied voltage, and the duration of the discharge.
2. The properties of the material are such that the stress within the material is sufficiently high to cause heating. Under these conditions a volume effect is in progress although a surface discharge is obtained.
3. The stress required to initiate an arc is higher than that which the material can withstand, and the failure occurs within the volume of the specimen.

For a given material, the type of failure depends upon the frequency of the applied voltage, the previous exposure history of the specimen, and the condition of the surrounding atmosphere during the test. The actual failure voltages are given in Tables III, IV, and V. Those values which represent partial or complete internal failures are designated accordingly. Only in the case of MFE and GSG is such extreme behavior observed. MFE fails internally at 100 mc at room condition and then at successively lower frequencies as moisture is absorbed. Ultimately, even the 60-cps test results in an internal failure. In the case of GSG, internal failures were observed from 2 to 100 mc after 6 months' exposure to humid atmosphere. Tests at high humidity indicate that films of water or discrete water droplets can cause momentarily low values of flashover strength. Except for these temporary effects, PF and PE maintained initial flashover strength throughout 6 months of exposure to all 9 conditions.

Table III shows the effects of temperature and frequency on flashover strength for the chosen materials over the com-

Table III. Flashover Strength, Kv, Initial Values

Temperature, C	60 Cps	1 Kc	2 Mc	18 Mc	100 Mc
MFE					
−55	10.6	10.1	8.2	9.0	7.0
25	9.9	9.8	8.0	8.0	6.3*
50	9.1	9.3	8.1	7.1	6.2*
85	8.6	8.5	6.7	6.3	6.3*
125	8.0	7.3	6.9	6.9	5.9*
GSG					
−55	12.0	12.0	11.1	11.7	9.0
25	10.8	10.6	9.8	9.4	8.7
50	9.5	10.3	8.7	8.1	8.5
85	9.6	9.6	8.5	8.1	8.0
125	9.8	9.2	7.5	7.5	7.5
PF					
−55	14.8	14.5	11.8	12.0	11.5
25	14.8	14.1	11.5	11.2	11.0
50	13.8	14.2	10.6	11.2	11.0
85	14.1	13.3	10.1	9.7	9.5
125	11.5	11.8	9.3	10.3	8.0
PE					
−55	14.9	14.9	10.7	12.5	11.5
25	15.0	14.9	11.3	11.0	11.0
50	13.2	14.2	10.0	8.8	10.7

*Flashover with attendant volume type failure.

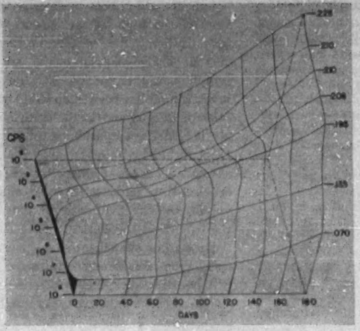


Fig. 11 (above). Effect of aging at 25C, 100 RH, on the dissipation factor of GSG

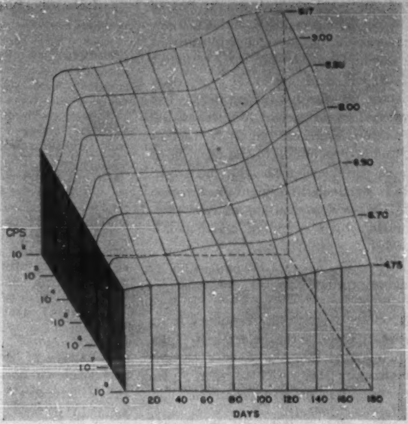


Fig. 12 (right). Effect of aging at 25C, 100 RH, on the dielectric constant of GSG

plete temperature range. It is noticeable that the flashover value decreases as the temperature is raised. However, the change is not as large or as definite as might be expected. On the basis of density, the breakdown voltage of air at −55C should be 1.82 times that at 125C. Since the flashover failure obeys no such rule, it is apparent that the properties of the solid materials are involved, even where the surface appears undamaged. In general, the variation of flashover value with temperature does not obey the simple density correction law. The data taken at −55C renders the highest flashover values for all materials, but definite conclusion is not possible because a slight frost on all parts proved unavoidable. Various experience with frost yields an impression that it is a rather good dielectric.

In most cases where a true flashover occurs, the effect of increasing frequency on flashover strength is not as drastic as it is on dielectric strength. The dielectric strength of a material can decrease by as much as a factor of 10 without seriously affecting the flashover strength. However, if the dielectric strength of the material decreases to the point where the solid becomes a limiting factor, then the flashover characteristics are seriously affected by increasing frequency. Under these conditions, the test actually becomes a parallel breakdown test.

The extent of the damage to the flashover sample depends on the duration of the discharge. Although this effect was not studied in detail it is worthy of note that only PF showed no change in flashover strength due to a prior test. A slight change in the texture of the surface is produced by the arc, which appeared to be the result of an erosion process, but this did not affect the flashover voltage of a succeeding test.

Table IV. MFE Flashover Strength, Kv, Exposure Data

Time	60 Cps	1 Kc	2 Mc	18 Mc	100 Mc
25C, 100 RH					
0	9.9	9.8	8.0	8.0	6.3*
1 hour	9.9	9.5	7.6	6.7	6.3*
1 day	9.9	9.8	7.2	6.0	5.0*
7 days	9.5	9.8			
30 days	8.8	8.4	5.0	4.9†	3.1†
180 days	2.0†	0.4†	0.7†	0.5†	0.5†
50C, 100 RH					
0	9.1	9.3	8.1	7.1	6.2*
1 hour	8.8	10.5	7.8	4.6*	5.0*
1 day	8.5	10.0	7.0	4.8*	4.5*
30 days	8.1	8.7	6.5	4.8*	0.5†
180 days	3.8†	3.3†	0.7†	0.7†	0.4†

*Flashover with attendant volume type failure.
†Internal volume type failure.

Table V. GSG Flashover Strength, Kv, Exposure Data

Time	60 Cps	1 Kc	2 Mc	18 Mc	100 Mc
25C, 100 RH					
0	10.8	10.6	9.8	9.4	8.7
1 hour	8.4	9.1	8.3	6.9	8.7
1 day	9.0	9.0	8.1	6.9	8.3
30 days	8.0	8.7	5.6	4.2*	0.8†
180 days	6.8	2.5†	0.8†	0.7*	0.7†
50C, 100 RH					
0	9.5	10.3	8.7	8.1	8.5
1 hour	8.7	8.2	9.4	7.5	8.0
1 day	8.6	8.1	9.3	7.4	8.3
30 days	7.3	8.8	7.6	7.0	6.6*
180 days	7.0	7.0	0.7†	0.7†	0.7†

*Flashover with attendant volume type failure.
†Internal volume type failure.

Moisture Absorption

Weight measurements made at regular intervals indicated that no change in weight occurred in the case of PF and PE at any of the conditions of the program. Both MFE and GSG absorb moisture and consequently either gain or lose weight, depending on the exposure condition. Table VI shows the percentage gain in weight for MFE and GSG at 25C, 100 RH and 50C, 100 RH. Table VII shows the percentage loss in weight for these two materials at several conditions. Since all samples were stored for many months at room condition (approximately 25C, 45 RH), the 25C, 50-RH and the 50C, 50-RH conditions had no definite effect on weight measurements. These conditions would have to be controlled with extreme precision in order to evaluate the small changes in weight which occur. At both of the 20-RH conditions MFE and GSG lose weight, indicating that some initial moisture is being dried out. The effect is more pronounced at 50 than at 25C. Weight measurements made on MFE and GSG after 2 months' storage at 25C in a sealed container with acti-

Table VI. Per-Cent Gain in Weight

Material	1 Day	10 Days	30 Days	90 Days	180 Days
25C, 100 RH					
MFE	0.01	0.07	0.18	0.51	0.82
GSG	0.04	0.07	0.14	0.17	0.29
50C, 100 RH					
MFE	0.01	0.13	0.33	0.78	0.98
GSG	0.05	0.09	0.22	0.25	0.27

vated silica gel showed weight loss of 0.084 per cent for MFE and 0.030 per cent for GSG. These values indicate that the weight losses at 25C, 20 RH and at 50C, 20 RH were due to the removal of moisture by the dry atmosphere. The greatest loss in weight occurs at the higher temperatures. At 125C part of the weight loss is due to the removal of volatile products from the materials. This is particularly evident in the case of MFE, which exhibits extreme discoloration by the end of the 6-month period.

Table VII. Per-Cent Loss in Weight

Material	1 Day	10 Days	30 Days	90 Days	180 Days
25C, 20 RH					
MFE	0.02	0.04	0.04	0.04	0.04
GSG	†	0.02	0.03	0.04	0.07
50C, 20 RH					
MFE	0.02	0.06	0.12	0.15	0.25
GSG	0.01	0.02	0.05	0.06	0.07
85C					
MFE	0.13	0.26	0.32	0.51	0.55
GSG	0.03	0.08	0.13	0.23	0.25
125C					
MFE	0.15	0.25	0.35	0.55	0.93
GSG	0.04	0.09	0.15	0.25	0.37

†No change in weight observed.

Conclusions

The effect of increasing frequency, from 60 cps to 100 mc, on the dielectric strength is much greater than the effect of varying temperature from −55 to 125C. In some cases an increase in temperature has no effect on dielectric strength, or may even cause an increase, but in no instance was such an effect observed for increasing frequency.

Although all materials have shown a marked decrease in dielectric strength with increasing frequency, in many cases flashover strength suffers only moderate diminution. Consequently such design considerations as the ratio of flashover voltage to puncture voltage change considerably with increasing frequency. Unfortunately, the relationship between these two types of failure would change with each differently shaped dielectric part or type of hardware, so a generally useful rule cannot be deduced from the simple structures which have been examined thus far.

Materials which absorb moisture suffer serious degradation of electrical properties when exposed to humid atmospheres. This effect, as borne out by the data on MFE and GSG, is much greater than that produced by temperature changes alone. It is interesting to note that the 60-cps dielectric strength test serves as a convenient and critical method of detecting moisture absorption by electrical means. The maximum effect is observed in the shortest time at this low frequency.

The primary purpose of this paper has been to present the data and descriptive comments which best summarize the results of the high-voltage aspects of the investigation. The results of the concurrent bridge measurements are presented in another paper.[5] At present the investigation has been extended to include other materials and in the complete analysis of all data an attempt will be made to confirm the proposal that dielectric strength is related to other electrical properties. Another current program involving the effects of natural exposure in the tropics on the electrical properties of ten materials should provide a means of correlating laboratory exposure data with actual field experience, thereby making the present data even more useful to the designer.

References

1. Dielectric Strength of Solid Insulation, **J. J. Chapman, L. J. Frisco.** *Electrical Manufacturing*, New York, N. Y., vol. 53, May 1954, pp. 136–43.

2. A Breakdown Cell for Measuring the Dielectric Strength of Solids at 100 Megacycles, **L. J. Frisco, J. J. Chapman.** *Review of Scientific Instruments*, New York, N. Y., vol. 25, Aug. 1954, pp. 733–37.

3. Behavior of Insulating Materials at Radio Frequencies, **J. J. Chapman, J. W. Dzimianski, C. F. Miller, R. K. Witt.** *Electrical Manufacturing*, New York, N. Y., vol. 48, July 1951, pp. 107–10.

4. Dielectric Breakdown Techniques, **J. J. Chapman, L. J. Frisco, J. S. Smith.** *The Electrochemical Journal*, New York, N. Y., vol. 102, Feb. 1955, pp. 67–72.

5. Surface and Volume Dielectric Losses, **J. J. Chapman, L. F. Blickley, E. A. Szymkowiak.** *AIEE Transactions*, vol. 74, pt. 1, July 1955, pp. 343–49.

◆

Discussion

C. N. Works (Westinghouse Electric Corporation, East Pittsburgh, Pa.): This study provides the user of solid insulation with much useful information. The use of three co-ordinates for the graphs, whereby dielectric strength is related to both time of aging and frequency, gives a striking picture of the dielectric performance of a material. Likewise, the graphs of dielectric strength, as a function of both temperature and frequency, give complete pictures of the performance of these materials. Previously such data had to be obtained from several sources. This information was usually difficult to correlate and often incomplete, particularly as regards dielectric strength versus frequency.

However, it should be noted that much information on dielectric strength as a function of temperature, especially for polyethylene, may be found in the literature, references to which are not cited in the paper. Since the paper deals with recessed electrodes, it is interesting to compare the results with the work done by Oakes[1] on very carefully prepared samples in which he measured the 50-cycle a-c dielectric strength as a function of thickness for thickness up to 15 mils at room temperature (20 C), and also the d-c and a-c dielectric strength as a function of temperature for thin samples. The dielectric strength of polyethylene is given in Fig. 8 as 1,091 volts per mil at 25 C and 60 cycles. The paper does not state whether the dielectric strength is given in rms or peak value. If the latter is used, the value of 1,091 volts per mil seems somewhat low. There is some question as to whether sprayed solder electrodes make good contact with the sample. The efficiency of these solder electrodes could be checked by the technique of Austen and Pelzer.[2]

In Fig. 8, the dielectric strength at 25 C decreases to only about 1/7 its value at 60 cycles as the frequency is increased to 100 mc. It is interesting to contrast this low value for about a 1-minute application of voltage at 100 mc to that obtained by Sucher, Farber, and Griemsmann[3] for microsecond pulses at 3,300 mc. Sucher and co-authors obtained a value of 12 kv per mil peak on 2-to-3-mil thick samples. This value is nearly identical with that obtained by Oakes[1] for a 1-minute application of 50-cycle voltage on samples of the same order of thickness. It is interesting to note that in both cases about the same total number of cycles of voltage were applied to the sample. It appears that each voltage maximum may cause some permanent damage.

The authors should state clearly in the text and on each graph whether the dielectric breakdown voltage is given in peak or rms values.

REFERENCES

1. The Intrinsic Electric Strength of Polythene and its Variation with Temperature, **W. G. Oakes.** *The Institution of Electrical Engineers*, London, England, vol. 95, pt. I, 1948, pp. 36–44.

2. Electric Strength of Paraffins and Some High Polymers, **A. E. W. Austen, H. Pelzer.** *Ibid.*, vol. 93, pt. I, 1946, p. 525.

3. Measurement of Breakdown of Polyethylene at 3,300 MC/Sec., **M. Sucher, H. Farber, J. W. E. Griemsmann.** Annual Report of the Conference on Electrical Insulation, National Research Council, Washington, D. C., 1952, pp. 16–17.

G. M. L. Sommerman (Westinghouse Electric Corporation, East Pittsburgh, Pa.): It is difficult to understand why the dielectric strength of such low-loss materials as polyethylene and polytetrafluoroethylene decreases inversely with frequency as much as is shown in Figs. 7 and 8. Even at 100 kc the dielectric strength is at least 40 per cent less than the 60-cps values at all test temperatures. Of course, the calculated dielectric loss at these high stresses is more than 1 watt per cubic centimeter, which gives some dielectric heating during the breakdown test. At 100 mc, the dielectric loss is still greater. On the other hand, the decrease in dielectric strength with increasing temperature is only slight in Figs. 7 and 8.

An alternative explanation is corona cutting from minute imperfections in the electrodes. Even though the authors have described additional precautions taken in their latest electrode designs, the possible effects of corona should be checked, because even a very small point can lead to high corona currents at these high frequencies. Perhaps this could be checked by testing a sample under high gas pressure. This would not eliminate electrode field-emission currents, but would greatly reduce ionization by collision in any gas bubble between electrode and sample.

It would help the reader to understand the phenomena involved if symbols were added to the dielectric strength charts to show the nature of the breakdowns at various temperatures and frequencies (as was done in the tables of flashover strengths).

G. S. Eager, Jr. (General Cable Corporation, Bayonne, N. J.): This useful paper, along with others published recently by Dr. Chapman and his co-workers, contains extremely valuable information concerning some insulations employed extensively on insulated conductors. We are particularly interested in the data presented on polyethylene since this material is being employed in increasingly large volume.

In the wire and cable industry polyethylene commonly is used in thicknesses from about 0.010 inch to about 0.250 inch at frequencies up to 1,000 mc, as well as at the 60-cps power frequency applications. The insulation most often is extruded over stranded conductors, which results, in general, in lower dielectric strengths than are obtained when smooth electrodes, e.g., a cable having a solid inner conductor, are used. The dielectric strength data presented by the authors are for a thickness of 0.030 inch; for greater thicknesses, approaching 0.250 inch the dielectric strength in volts per mil would be materially lower. Do the authors have available dielectric strength values of polyethylene in greater thicknesses than 0.030 inch in the frequency range studied?

The authors do not mention the specific type of polyethylene employed in their tests; we assume it to be nonpigmented heat-stabilized polyethylene to military specifications (see reference 1 of this discussion). Since the wire and cable industry regularly uses several types of polyethylene in various colors and since some electrical properties are affected by these colors, etc., it is felt that the value of this paper would be increased if the authors would define specifically the type of polyethylene employed in their tests.

REFERENCE

1. DIELECTRIC MATERIAL, POLYETHYLENE. *Military Specification MIL-D-3054A*, U. S. Government Printing Office, Washington, D. C., July 11, 1952.

R. W. Atkinson (General Cable Corporation, Bayonne, N. J.): It seems useful to discuss some relationships between these measurements and certain pertinent fundamental electrical properties. Probably the only electrical property of the solid dielectric which affects the flashover voltages is the dielectric constant. The stress in the air surrounding the solid dielectric increases with increasing dielectric constant; thus the flashover voltage will decrease with increasing dielectric constant of the solid. This accords also with experience and, as I see it, with the authors' data. Where tracking and burning takes place, the resultant change in the properties of the surface will affect the flashover, but the authors have pointed out that, on these materials, there has been little tracking and burning until actual flashover has occurred.

I was inclined initially to attribute the decrease which the authors have found in breakdown strength with increasing frequency chiefly to dielectric loss. Clearly, from their description of the failures and from the dissipation factors given, this is true in some cases, perhaps for all of the tests on the two composite insulating materials.

However, upon re-examination of the authors' data, and in the light of their oral discussion, it must be conceded that few if any of the breakdown tests on the teflon and polythene can be explained on the basis of dielectric heating. Thus, if we assume the same dissipation factor at 100 as at 18 mc, the loss at 100 mc would be $5^{1}/_{2}$ times as great at a given voltage as at 18 mc. Therefore, to give the same loss and heating at the higher frequency the voltage could be only 43 per cent of that at the lower frequency (noting that the loss varies as the square of the voltage). In the case of teflon the breakdown voltage is two-thirds as much at the higher frequency and the ratio is even higher for polyethylene. From this type of analysis it seems that the failures on these materials are not due to dielectric loss at 18 mc or under. With the amount of loss which must occur at the stresses shown and with dissipation factors which occur on these materials, both of these must have been approaching thermal failure at the 100-mc frequency, but there is no direct evidence here that the failure was thermal on these materials even at 100 mc.

Because the large effect of frequency which the authors have found upon the breakdown strength of these materials cannot be explained on a simple thermal basis, their data are of much interest and importance. As these authors found in connection with their flashover tests, both PF and PE are subject to a kind of erosion when subjected to surface discharge. If, within the stressed region, bubbles appear of sufficient size (probably a diameter at least as great as 2 or 3 mils would be required) to suffer ionization, then the failure might be explainable on the basis of erosion within these bubbles and the resulting rapid growth of the spaces at the rapid repetition rate of discharge which would occur at the higher frequencies employed.

With thermal causes eliminated and if some such explanation as just suggested is not valid, it would seem that the explanation must be in terms of molecular properties. Electric stresses do produce mechanical stresses, and do produce some molecular displacements. For large mechanical displacements, these materials are imperfectly elastic. Perhaps some permanent displacement in the structure of the dielectric takes place at each cycle, which is multiplied in the successive thousands of millions of cycles, leading eventually to rupture. A satisfying evaluation of the causes of this effect of frequency upon breakdown strength may be an important step in understanding dielectric behavior of materials.

J. J. Chapman, L. J. Frisco, and J. S. Smith: With respect to Mr. Works' comments, it must be mentioned that rms values of voltage are used throughout the paper. The work of Oakes (see reference 1 of Mr. Works' discussion) indicates that the dielectric strength of PE is much higher at direct current than at 50 cps except for quite thin films. It also appears evident that 50 cps values are higher than 60 cps values. If the curve of Oakes were extrapolated to 30 mils, it would appear that our 60 cps value might be 30 per cent lower than his 50-cps value.

This degree of confirmation is not very surprising because our results have been checked by several arrangements, one of which greatly resembles that of Austen and Pelzer (see reference 2 of Mr. Works' discussion). Brass electrodes have been encapsulated in PE and values of breakdown have checked with those obtained with cavities coated with solder or conducting silver paint. The technique of Austen and Pelzer is useful with direct current or low frequency; however, it would not serve for wide frequency range exploration, if graphite coatings are essential to the method. The introduction of complex composites and assemblies around the high-voltage electrode leads only to surface heating and possible gas evolution in the immersion medium at high frequency. The early work of Becker[1] is of interest in the discussion of electrodes.

Corona arising within the volume of a dielectric is only a function of its interstitial structure and may be properly considered a property of a given material. Mr. Sommerman raises questions regarding external corona that might be introduced as a result of faulty application of electrodes. Imperfection, in the general contour, of all electrode designs has been eliminated from our tests by equivalent tests in air and oil. At small spacings (say 30 mils) our electrode elements are required to give results in air and oil closely agreeing with 1-inch diameter spheres; this eliminates the question of significant field divergence. Imperfections in the surface finish of a cavity, which could still cause corona even though the general contour is good, have been considered. In the case of many materials it has been possible to drill, grind, and mold cavities for recessed electrodes. In addition, sprayed solder and conducting silver paint have been used to coat cavities. Thus with a variety of electrode sizes, coatings, and methods of making cavities, applied among those materials amenable to different approaches, general agreement prevails. The mechanical handling and surface application at the point of engagement of the high-voltage electrode is more pertinent than considerations of contour. Naturally, the authors feel that corona introduced as a result of faulty treatment of the surface of a material has been avoided because of the agreement among different methods. It must be added that different liquid immersion media have been used without changing the results.

Unfortunately, there is no certain method of proving that a surface is without imperfection or bubble; this uncertainty holds for any method and is not limited to the ones used in this investigation. However, until higher values of voltage can be obtained with the same thickness and under the same test conditions, present phenomena may be ascribed to effects other than bubbles. In our work, the breakdown level of various insulating materials has increased annually, indicating improvement in test methods rather than improvement in materials. However, such raising of breakdown level with respect to frequency was slight during the past year; apparently for some time exceptionally low values that drastically affect the average result have been absent. It is felt that, for these particular test circumstances, the ultimate strength of the materials has been reached. Spread in results may in time prove to be inherent to certain molecular distributions. No attempt has been made to make a statistical study of spread as a function of frequency. The data did not quite reach statistical proportions. The results are considered reliable to ±15 per cent.

Unfortunately, the breakdown versus thickness data which Mr. Eager refers to cannot be provided at present. However, odd bits of data at 60 cps on different lots of PE indicate that the difference in dielectric strength between 30 and 60 mils is not great; in fact, the spread in results outweighs any difference at those thicknesses. At thicknesses of less than 30 mils, however, the rise in volts per mil with further decrease in thickness appears to be most emphatic. It is hoped that in time our program will include data at various thicknesses over the entire frequency range. At 60 cps, a properly made specimen of PE of 250 mils' thickness might be expected to require 200 kv for breakdown. Such a voltage is prohibitive in our general planning of tests. The PE which was used in the breakdown program was unpigmented, stabilized, 1/2-inch sheet stock.

Mr. Atkinson expresses interest in the flashover data as well as the puncture data. This is gratifying because little was said in the paper beyond the presentation of flashover results. Yet all industrial designers realize that flashover in some apparatus is far more probable than the puncture type of failure. The information presented is intended to be useful, but far more is required before generalizations can be drawn. In studying this surface property, spurious effects of very subtle nature may eliminate all finesse from the results, i.e., handbook data may be accumulated but no law evolved. For example, it has been found that materials such as PE or PF may readily acquire extreme static charge; in such cases the flashover value may be reduced to one-third of its normal value in an alternating field. With such materials, complete freedom from the effects of static electricity proves to be tedious to achieve for the purposes of experiment. For the designer, the problem of safeguarding against such spurious effects is even more troublesome.

Regarding the puncture data, there seem to be two phases of the discussion: first, the validity of the methods and, second, the explanation of the results. Aside from the details presented in the paper and in the discussion, it must be pointed out that these methods were tried on many materials. Further, experience with each material brought out some small detail that was useful in the case of other materials. Drastic decrease in dielectric strength as a function of increasing frequency appears to be true in the case of all solid insulation.

Although the generation of heat within a specimen may improperly conclude an otherwise purely electrical breakdown test, there are materials and frequency regions where heat merely appears to accompany breakdown as a moderate influence. In such cases heat appears to play no directly causative role, it is merely a measurable effect, indicating that disorderly motion of electric charge occurred under the action of electric stress. Temperature rise indicates a degree of randomness within the dielectric during the application of electric stress, which is perhaps a factor to be considered in theories of ultimate stress based upon mechanistic concepts. Perturbations generated by the alternating electric field may create ionization at random points within the dielectric and greatly affect the ultimate strength as a function of frequency.

Reference

1. Messung von Durchschlagsfeldstärken fester Isolierstoffe im Frequenzbereich 1 MHz bis 15 MHz, R. Becker. *Archiv für Electrotechnik*, Berlin, Germany, vol. 30, 1936, p. 411.

A Variable Frequency Magnetic-Coupled Multivibrator

R. L. VAN ALLEN
ASSOCIATE MEMBER AIEE

IN THE last decade, the technology of magnetic core materials has advanced such that magnetic cores are being used in many electronic applications over and above conventional use in transformers. Take, for example, magnetic amplifiers, computers, television components, to mention a few.

The particular application of magnetic cores discussed in this paper is their use in obtaining multivibrator action of a simple circuit which has the advantage of being easily controlled in frequency over several octaves. This magnetic-coupled multivibrator can be a power device capable of delivering watts of power and yet be frequency controlled over several octaves with only a few milliamperes of control current. Although the frequency range of a particular unit is limited, three experimental models have been built collectively covering a frequency range of 200 cycles per second to 700 kc. The multivibrator action of the circuit being basically a switching process permits it to be readily adapted to transistors or vacuum tubes with the switching rate controlled via the magnetic core material. The circuit performs particularly well with transistors since a low impedance switching device is an essential component of the circuit required to realize high efficiencies.

This circuit should prove to be of value in many conventional multivibrator applications, especially where appreciable power- or frequency-modulated output is desired. Also, it should find use in applications where multivibrators require frequency stabilization against drift; the fact that the frequency can be readily controlled facilitates frequency drift compensation.

Background

The oscillator described in this paper bears a resemblance to the conventional free-running multivibrator both in the wave forms produced and in the theory of operation. Bistable states of gating and

---

Paper 55-75, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing November 19, 1954.

R. L. Van Allen is with the Naval Research Laboratory, Washington, D. C.

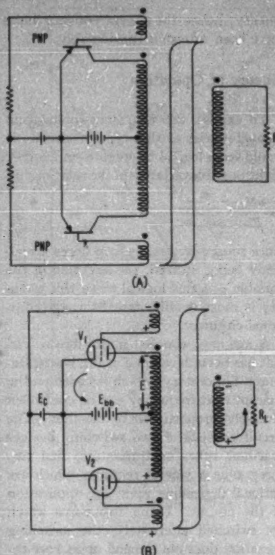


Fig. 1(A). Royer transistor circuit. Fig. 1(B). Equivalent vacuum-tube circuit

cutoff produce essentially rectangular wave form in both devices. Examination reveals that a closer similarity exists in that the stable oscillating frequency of each device depends upon a direct integrating process. The frequency-determining components of the conventional multivibrator are the well-known *R* and *C* with their associated time constants. If the charge *Q* on the condenser is considered where

$$Q=\int i\,dt$$

the time dependent *Q* which, after all, determines when the multivibrator will fire is by definition the time integral of current. Now, an analogous multivibrator circuit is arranged where, instead of applying an *RC* network, magnetic circuits are utilized with core material having a high degree of remanence such that the time-dependent quantity be the integral of voltage instead of current. It follows that an analogous multivibrator could be constructed having frequency dependent upon the time integral of voltage. For this circuit configuration, consider the flux in the core where total change in flux

$$\phi=\frac{1}{N}\int e\,dt$$

for constant voltage

$$\phi=\frac{E}{N}t \text{ volt-seconds}$$

$$t=\frac{\phi}{E}N$$

Then, time is directly proportional to the flux ϕ and inversely proportional to the voltage *E*. This concept has been demonstrated in a circuit disclosed by Royer in an unpublished AIEE paper of the Summer and Pacific General Meeting, June 1954, which could be described as a magnetic-coupled multivibrator. This circuit, as shown in Fig. 1(A), was designed for use with transistors, and exhibits properties which make it especially suited for the low-impedance characteristics of such devices. An equivalent vacuum-tube circuit is shown in Fig. 1(B), and, for purposes of simplicity, the func-

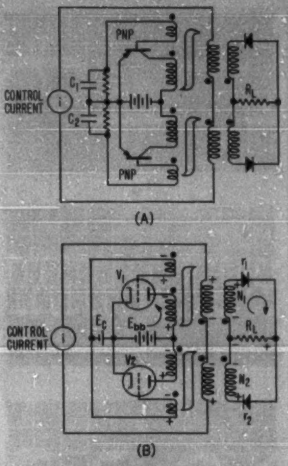


Fig. 3. Variable frequency magnetic-coupled multivibrator circuit using (A) transistors and (B) vacuum tubes

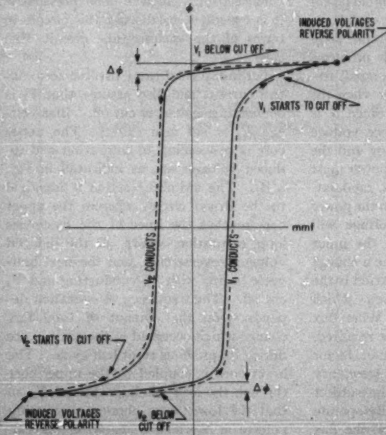


Fig. 2 (left). Operation about hysteresis loop

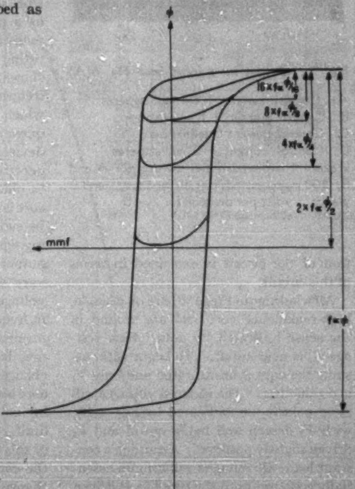


Fig. 4 (right). Operation about minor hysteresis loops

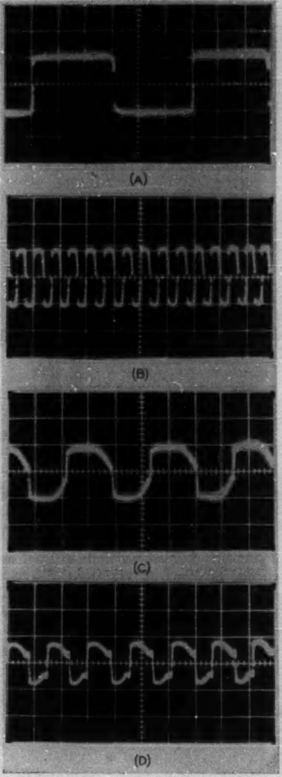


Fig. 5. Output wave forms from Fig. 6(A)

A: Control current +1.4 milliamperes
B: Control current +1 milliampere
C: Control current 0 milliampere
D: Control current —2 milliamperes
x axis: 1 millisecond per division for A and B; 100 microseconds per division for C and D
y axis: 5 volts per division for A—D

tion of the circuit is explained in terms of this circuit.

All windings in Fig. 1(B) are on a single, high-remanence core and are wound in the sense indicated by dots. Bias voltage E_c is near cutoff. To begin with, assume the core is unsaturated and only V_1 is conducting. The induced voltages will have polarity as indicated in Fig. 1(B), with V_2 driven well below cutoff and V_1 driven slightly positive. Assuming a constant load, all voltages will remain essentially constant until the core flux is driven to saturation. At core saturation, the rate of change of flux, which up until now has been relatively constant, will suddenly be reduced causing a reduction in all the induced voltages. The sharpness of this voltage change will depend upon the squareness of the hysteresis loop. The decrease in induced voltage allows the grid of V_1 to return toward cutoff which, in turn, causes the plate current to decrease and starts a reverse flux change toward the residual flux level of the core. This small reverse change in flux $\Delta\phi$, shown in Fig. 2, is the key to the multivibrator action of this circuit. Because of this small inductive kick, the polarity of all induced voltages will be reversed instantaneously, causing V_1 to be driven well into cutoff and, at the same time, raising the grid of V_2 to conduction. With V_2 conducting, the core flux changes toward the lower knee at a rate determined by the voltage E_{bb}. Thus, the time required to complete 1/2 cycle will be

$$t=\frac{\phi}{E}N$$

and, as Royer states, the frequency is indeed proportional to the d-c input voltage. This circuit serves the field of magnetic amplifiers very well, where there has been a long-felt need for low-impedance high-efficiency a-c power supplies, especially in the upper audio frequency range, and represents a tremendous contribution to the field. The transistor version of this circuit is quite remarkable, since it is a switching type of circuit and, as a power supply, may transfer power many times in excess of the nominal power rating of the transistor. The circuit, as shown, does have limitations for large frequency variations as well as problems which may be encountered in small frequency variations, particularly when the device is delivering power. Since it is necessary to modulate the core voltage to obtain a change in frequency and the core is delivering load current, power must be supplied from any external, modulating signal source proportional to the power output. Also, the output voltage will vary with frequency because the input voltage was changed to produce a change in frequency. A further limitation is the maximum operating frequency which may be obtained practically. When flux changes from knee to knee are required, core losses will be the predominant factor in limiting a practical upper frequency limit. The circuit which is the subject of this paper does not necessarily operate the cores knee to knee, in fact, the flux change is smaller with increased frequency; hence, the upper frequency limits have been extended considerably.

Theory of Operation

To exploit the frequency-modulation characteristics of this type of circuit, it would seem logical to operate on another of the possible variables in the relationship

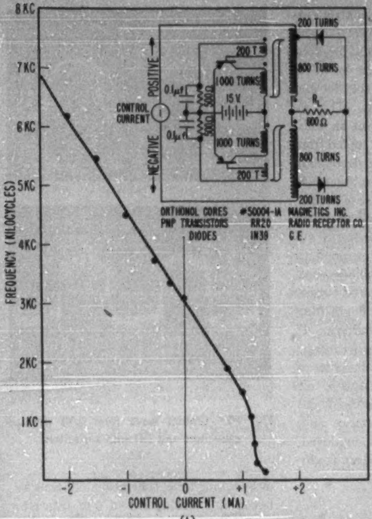
$$t=\frac{\phi N}{E}$$

Since square hysteresis loop cores are already being utilized, the selection of the variable ϕ seems logical since this quantity is most readily varied through additional circuitry.

A magnetic-coupled multivibrator circuit has been developed at the Naval Research Laboratory, which is based on the concept that frequency can be controlled inversely proportional to the flux. The circuit consists of two switching devices (vacuum tubes or transistors) and two cores, plus a pair of rectifiers which are optional depending upon the application of the circuit. Using two cores which are switched alternately, the switching operation does not depend upon knee-to-knee saturation to complete the cycle. Instead, each core need saturate at only one knee and need reset only part way on alternate half-cycles. Then, for each half-cycle, only one core proceeds to saturation, permitting the other core to be reset a predetermined amount in a manner similar to push-pull magnetic amplifiers.[1,2,3] The magnetic-coupled multivibrator is shown in Figs. 3(A) and 3(B) in both the transistor and vacuum-tube configuration. As was done previously, it is expedient to discuss the circuit in terms of the vacuum-tube circuit, Fig. 3(B).

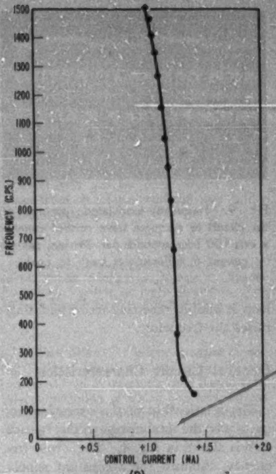
For initial conditions, assume zero control current and also assume that V_1 is conducting and V_2 is cut off. Bias voltage E_c is set near cutoff. The upper core is proceeding to saturation and induced voltages are as indicated in Fig. 3(B). The switching action is identical to the Royer circuit wherein the upper core reaches the knee of the hysteresis loop, saturation occurs, all the induced voltages reverse sign, and the next half-cycle begins with V_2 conducting and V_1 cut off. The frequency of operation depends upon the amount of total flux change which occurred in the gating core during its previous reset half-cycle. The lower core is coupled to the upper core through the load windings N_1 and N_2, so that the lower core obtains a resetting voltage on N_2 which is somewhat less than that appearing on the upper core wind-

Vol
Pt.
195



(A)

Fig. 6(A) (left). Control current versus frequency characteristics

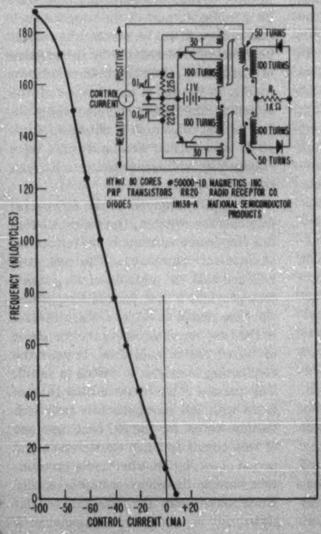


(B)

Fig. 6(B) (right). Lower frequency portion of curve in A

ing N_1.

Because of winding resistance and forward rectifier drop, the resetting core is never reset completely by the firing core. Therefore, the voltage drop around the loop, N_1, r_1, r_2, and N_2, will determine to some extent the zero control current frequency. Since the resetting voltage will be decreased by the amount of this drop, the flux level to which the resetting core is driven will determine the flux change to occur when the core fires on the next half-cycle. Then, for a constant supply voltage E_{bb}, the frequency may be increased or decreased depending upon the direction and magnitude of control current applied. For example, in Fig. 3(B), if the signal current is counterclockwise in the loop, it will aid the resetting action and, thus, produce more resetting flux which lowers the frequency or, if the signal passes clockwise through the



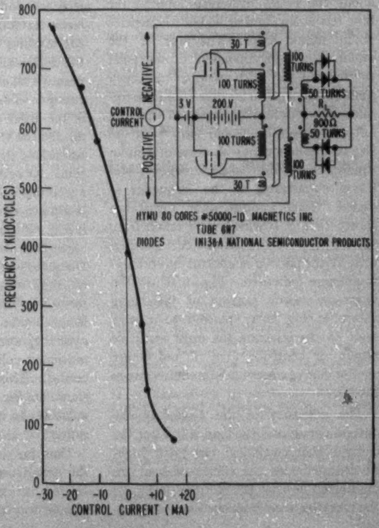


Fig. 7 (left) and Fig. 8 (right). Control current versus frequency characteristics

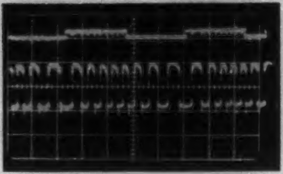


Fig. 9. Frequency-modulated response of the circuit to a square wave control current; x axis 100 microseconds per division, control current 5 milliamperes peak to peak

loop, it inhibits resetting action and, thus, raises the frequency.

General Circuit Characteristics

The relationship of the operating frequency to the total change of flux in each core is shown in Fig. 4. The lowest frequency obtained for a constant supply voltage E_{bb} and a fixed number of primary turns will depend upon the knee-to-knee flux change. This lower frequency may be defined as the base frequency and, in Fig. 4, this is the operating frequency f. Higher operating frequencies such as $2f$, $4f$, $8f$, etc., are established by the control current acting as an magnetomotive force bias which determines the amount of flux reset in a given core and thus establishes the frequency of operation. In other words, the cores are operated about the entire hysteresis loop at the base frequency, and at higher frequencies about minor loops. The size of the minor loop being inversely proportional to the frequency with an upper frequency limit of about 20 to 30 times the base frequency. It would be expected that the output wave form would vary with the size of the minor loop showing maximum squareness at the base frequency f and losing squareness as the frequency is raised. Some typical wave forms are shown in Fig. 5 which were taken from the frequency characteristic curve of Fig. 6(A) at the particular control current and frequency indicated. Fig. 6(B) is the lower frequency portion of the curve shown in Fig. 6(A), plotted to a large scale to demonstrate the most sensitive region of control. Figs. 7 and 8 are similar curves covering different frequency ranges.

The efficiency of the diodes in the output circuit and the load will affect the control characteristics, but with high-efficiency diodes the effects of load are minimized because the rectifier decouples the resetting core from the load.[4] Referring to Fig. 3(B), assuming control current in a counter-clockwise direction and r_1 and r_2 high-efficiency diodes, the voltage across the windings of the lower resetting core is negative at the dot and blocked by r_2 from the load. This allows the lower core to be reset an additional amount and yet be decoupled from the load resistance R_L. The decoupling provided by the diodes is also important in terms of the time of response of the circuit to changes in frequency. When rectifiers having poor back resistance are used, the circuit is slow to respond to signals which change the oscillating frequency. A frequency-modulated output is shown in Fig. 9 where the modulating signal is a square wave of current. It can be seen that the change in frequency is not instantaneous being of the order of several cycles. The rectifiers used in this case were *1N138-A* silicon junction diodes having very high back resistance and, when less efficient diodes were employed, the time of response for frequency modulation was longer.

Transistor Characteristics

The transistor circuit of Fig. 3(A) in principle operates the same as the vacuum-tube circuit previously described. Its operation, however, does differ in that no bias voltage is necessary and its efficiency is higher. The PNP junction transistors in this circuit are connected with a common emitter and are switched from low collector resistance to high collector resistance by voltage appearing on the winding connected to the base. In the conducting state, the base is more negative than both the collector and the emitter. The collector resistance is low, of the order of several ohms, permitting an efficient transfer of the supply voltage to the load as the core proceeds to saturation. At saturation, when the base voltage drops to nearly zero, the collector resistance rises further so that at cutoff it may be 100,000 ohms or more, depending on the transistor characteristics. The transistors in either state of conduction or cutoff dissipate very little energy internally which permits raising the collector conduction current many times over the normal rated value as long as maximum collector dissipation is not exceeded. This aspect of Royer's circuit, shown in Fig. 1(A), makes it especially valuable as a magnetic-amplifier a-c power supply.

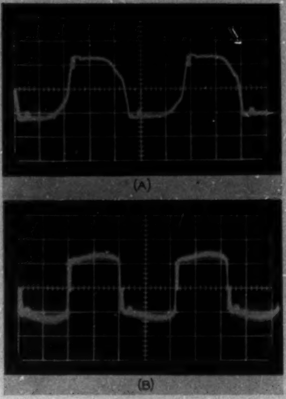


Fig. 10. Output wave form (A) without capacitors and (B) with capacitors

Thus far, only the bistable states of the transistor circuit have been considered, i.e., conduction and cutoff. An equally important consideration is that of switching time, specifically the time between full conduction of one transistor until full conduction of the other. When one type of 500-milliwatt power transistor was used in either the circuit of Figs. 1(A) or 3(A), the switching time became an appreciable percentage of the period of each cycle, as shown in Fig. 10(A). This, of course, contributed to the internal dissipation of the transistor for, to use the transistors most efficiently in these circuits, the switching time must be a negligible part of the time for each half-cycle. It was found that the switching time could be reduced by the addition of capacitors C_1 and C_2 in the circuit shown in Fig. 3(A), and a comparison of the switching time before and after the addition of the capacitors is shown in Figs. 10(A) and 10(B). Similar results were obtained by adding capacitors to Royer's transistor circuit.

A complete analysis of the factors influencing the switching time of transistors carrying heavy currents bears further investigation. However, in this case, it is believed that the addition of the capacitors provides a low impedance path in the base circuit to allow a quick reverse in the base current clearing the transistor of stored charge and, thus, hastens the conducting transistor to switch to cutoff. This parallel R and C should not be confused with self-biasing circuits used with vacuum tubes, for the RC time constant of this circuit is much shorter than the period of oscillation, whereas for vacuum-tube biasing, the time constant is usually much longer than the period of oscillation.

Summary

The magnetic-coupled multivibrator presented in this paper offers the advantage of delivering power and, at the same time, is capable of being frequency controlled over several octaves with small alternating or direct currents. It may be constructed by using vacuum tubes or transistors. The convenience of having continuous frequency control should prove of value in multivibrator applications where frequency stabilization and control are required as well as opening new fields of application for frequency-modulated transistor oscillators.

References

1. On the Mechanics of Magnetic Amplifiers, R. A. Ramey. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

2. The Influence of Magnetic Amplifier Circuitry Upon the Operating Hysteresis Loops, H. W. Lord. *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 721–28.

3. Self-Saturation in Magnetic Amplifiers, W. J. Dornhoefer. *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 835–50.

4. The Effect of Core Material on Magnetic Amplifier Circuits, L. J. Johnson. *AIEE Transactions*, vol. 71, pt. I, Jan. 1952, pp. 26–31.

◆

Discussion

George C. Uchrin (Signal Corps Engineering Laboratories, Fort Monmouth, N. J.): The Signal Corps Engineering Laboratories have been conducting a study of a push-pull self-excited square-wave power oscillator for use in a power conversion system which converts low-voltage direct current to high-voltage direct current for the operation of various types of communication equipments. Frequency of operation and a means control have been investigated to determine the effect of the oscillator upon the over-all efficiency and size of the converter. A method of frequency control which gave results similar to what the author has obtained was investigated. It consisted of placing a variable d-c bias current in a separate winding on the core. This method afforded a 5-to-1 control in the frequency.

Core materials other than those having a rectangular hysteresis loop have been used in closed circuit and air-gap transformers with good results. Attention is invited to an article[1] showing wave forms which are almost ideal, and which were obtained with just the basic circuit of the square-wave oscillator without any additional components.

Reference

1. A New Self-Excited Square-Wave Transistor Power Oscillator, George C. Uchrin. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 43, Jan. 1955, p. 9.

R. L. Van Allen: It is encouraging to see the interest that is developing in the combined field of magnetics and transistors, the compatibility of which two components has extended tremendously the horizons for applications of magnetic circuits.

A Continuous-Flow Moisture Detector

M. F. BEAVERS
MEMBER AIEE

E. J. SHIMANSKI
NONMEMBER AIEE

E. F. TIMPANE
NONMEMBER AIEE

MOISTURE is one of the main foes of electric apparatus. In liquid-filled apparatus such as transformers, excessive moisture in the liquid can be a reflection of a similar situation in the solid insulation. The problem of detecting accurately these minute quantities of water in the liquid is an extremely important one and has received considerable attention in the past.

In 1935, Karl Fischer reported a reagent containing iodine and sulphur dioxide which could be used for the detection of part per million quantities of moisture in organic and inorganic materials.[1] Initially, the titration was carried out using a visual end point which was often difficult to detect. In 1940, Almy, Griffin, and Wilcox described an electrometric method for determining the end point which involved a potentiometric technique.[2] In 1946, Cornish devised a circuit suitable for direct titration with Karl Fischer's reagent using the so-called dead-stop technique.[3] A modification of this circuit has been used by the authors for several years and has been found to give very precise results.

This progress has resulted in a relatively quick and accurate method for the determination of water in insulating liquids. Since the use of this method requires procurement of a representative sample and transportation of this sample to the titration assembly, it is at times both inconvenient and inadequate. To interpret the results correctly, it is essential that no change in water content occur during the transport of the sample. Extreme care must be exercised in this sampling procedure to prevent addition or loss of moisture during the operation. In many instances such a transfer is essential because of the nature of the system. However, for working oil lines, filtering operations, or other installations where flowing oil is present, a distinct advantage would be obtained by the use of a detector which would continuously indicate the amount of moisture in the oil. This is particularly true when the water content of the oil is changing. The problem of transfer and multiple analyses is completely eliminated.

In this paper, an instrument is described which when placed in a flowing oil system, either directly or as part of a by-pass, is capable of indicating the water content of the oils by means of changing electrical properties. Other devices employing electrical means for indicating water content have been described in the literature and in patents.[4–9] These devices are adaptable to certain applications but none is suitable for detection of minute quantities of water in a flowing unexposed system.

The detector described in this paper was developed primarily for use in transformer installations. It can be connected to factory supply lines to permit periodic checking of moisture content or as part of a signaling device to warn of excessive moisture in the oil. In filtering operations this device, when installed on the outlet line, can be used to monitor the water content of the oil. Thus, it will readily indicate when the filter elements should be changed and when the oil has been dried to the desired degree.

Description of the Moisture Detector

The moisture detector consists essentially of a moisture-sensitive fabric sandwiched between two metal plates with perforations. The liquid under investi-

---

Paper 55-83, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N.Y., January 31–February 4, 1955. Manuscript submitted October 21, 1954; made available for printing November 19, 1954.

M. F. Beavers, E. J. Shimanski, and E. F. Timpane are with the General Electric Company, Pittsfield, Mass.

The authors gratefully acknowledge the contributions and sustained interest and encouragement throughout the course of this work of B. D. Treanor, M. I. Alimanski, E. L. Raab, and K. W. MacFadyen. We are grateful to Paul E. Flanagan for his analytical work and many helpful suggestions which contributed to the success of this investigation.

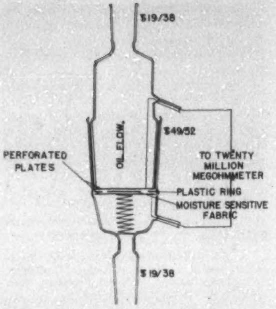


Fig. 1. Diagrammatic drawing of continuous-flow moisture detector

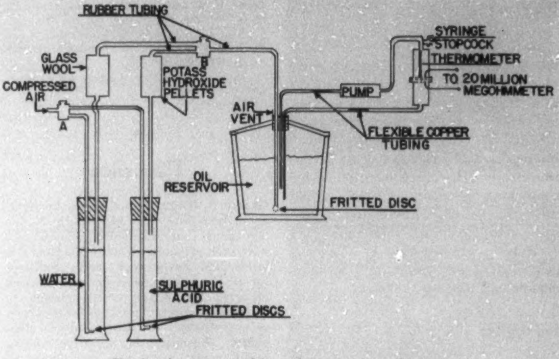


Fig. 2. Apparatus used for evaluating moisture detector

gation is passed through this element and moisture from the liquid is adsorbed by the fabric in an amount proportional to the relative moisture content of the liquid. This varying amount of moisture in the fabric causes changes in the electrical resistance between the two perforated plates. Thus, by measuring this resistance, a measure of the relative moisture content of the liquid is obtained.

A major part of the investigation was carried out with a laboratory-constructed moisture detector, illustrated diagrammatically in Fig. 1. The essential glass parts were made of the two components of a *49/52* standard taper ground glass joint. The upper electrode was held in place by the male member of the joint while the lower one was held in place by a spring. Between the perforated electrodes was a circular sheet of a moisture-sensitive fabric surrounded by an annular ring of polytetrafluoroethylene. This particular plastic was chosen as a spacing material because of its high insulation resistance and low moisture adsorbency. Because of the fragility of glass, certain portions of the investigation were made by using a model made of methyl methacrylate plastic having threaded ends to accommodate pipe fittings.

Experimental Procedure

To evaluate the detector, the arrangement illustrated schematically in Fig. 2 was set up. It consisted essentially of a reservoir for the oil, a means of adjusting the water content of the oil, the moisture detector, a centrifugal pump to circulate the oil through the cell, a means of measuring the resistance across the electrodes of the detector, and a valve for sampling the oil.

To increase the water content of the oil in the reservoir, stopcock *A* was adjusted to supply air to the fritted glass disk situated near the bottom of the water tower. The disk produces widely dispersed air bubbles which rise through the water and become saturated with moisture. The air then passes through a glass wool tower which removes entrained water droplets. The moisture-laden air then goes through stopcock *B* into the fritted glass disk at the bottom of the oil reservoir. The finely dispersed air bubbles pass up through the oil and out the air vent at the top of the reservoir.

The procedure for drying the oil was identical with the foregoing except stopcock *A* was adjusted to allow the air to pass into the sulphuric acid tower which thoroughly dried the air. A potassium hydroxide tower removed any acid which might have been carried by the air stream.

During the course of wetting or drying the oil, the centrifugal pump circulated the oil from the reservoir through the moisture detector. The resistance between the perforated electrodes was obtained by using an Electronics Instruments Ltd. 20,000,000-megohmmeter. When the resistance reached the desired range, the scrubbing was stopped and the air vent at the top of the reservoir was closed. When equilibrium between the water in the oil and the water in the moisture-sensitive fabric was reached, as indicated by a constant resistance reading over a period of 5 to 10 minutes, the temperature was recorded and a sample taken into a 30-cubic-centimeter medical syringe through a special stopcock designed for this purpose and soldered to the copper tubing. The water content of this oil was determined by using an electrometric Karl Fischer titration. From a knowledge of the saturation water content of the oil at the same temperature, the per-cent saturation was calculated.

Results and Discussion

General

The operation of the detector depends primarily on the rapid attainment of equilibrium between the water in the oil and the water in the fabric. To provide the quickest possible response, the instrument was designed to allow the oil to pass directly through a very thin, high-

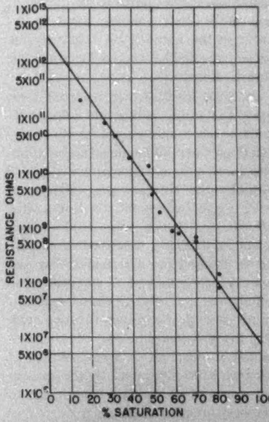


Fig. 3. Log of resistance versus per-cent saturation of water in transformer oil, 0.003-inch fabric with 0.010-inch spacer

Vol. 7
Pt. 1
1955

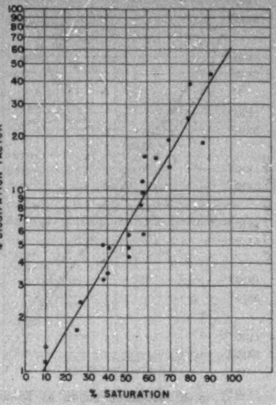


Fig. 4. Log of dissipation factor at 1 kc versus per-cent saturation of water in transformer oil, 0.003-inch fabric

porosity fabric. For an extreme change in relative moisture content, only 2 gallons of oil were required to reach the new equilibrium. During the course of the investigation two designs of the detector were used. The first one was constructed of two sections of plastic between which the electrodes, spacer, and fabric were placed. The other was a glass cell constructed as previously described. Most of the fundamental studies were carried out using the glass design. Results obtained with the plastic cell are reported later in the paper.

Fundamental Studies

The first stages of the investigation were carried out to obtain data using various electrode arrangements. Preliminary studies indicated that plotting the water contents of the oils in terms of per-cent saturation gave relationships which were practically independent of temperature and oil quality. The advantages of this type of graph over one involving absolute water contents are explained more fully later. The resistance values were plotted on a log scale. Fig. 3 shows the straight-line relationship obtained using an 0.003-inch circular sheet of moisture-sensitive fabric surrounded by an annular ring spacer 0.010-inch thick. All but one of the points are within 5 per cent of the line over the entire range of per-cent saturation. Since the oil, at saturation, had a water content of about 60 parts per million, the maximum deviation was approximately 3 parts per million, which is sufficiently accurate for most applications. The detector was also evaluated using an 0.005-inch spacer with the 0.003-inch fabric. The results gave a fairly good straight line having a similar slope but displaced toward lower resistance values. The spread in the data was slightly greater than was obtained with the 0.010-inch spacing. Elimination of the spacer and utilization of a larger diameter sheet of fabric also gave more erratic results but, generally, followed the straight-line trend. The line obtained from these data was removed even further toward lower resistance values as would be predicted from the smaller separation of electrodes and larger fabric area. In general, the smaller spacings seemed to give more erratic results. It was felt that smaller gaps were more greatly affected by minute particles in the oil.

Much of the information contained in patent and technical literature concerning moisture-detecting devices has been obtained using power-factor measurements as a means of estimating water content. For this reason, a brief study of this relationship, using the continuous-flow moisture detector, was carried out. The experimental techniques were identical with those used for resistance measurements except that a General Radio type-*716-c* capacitance bridge was used in place of the megohmmeter. Dissipation factor and capacitance measurements were made at 1, 10, and 100 kc. Sixty-cycle measurements were very erratic and could not be used. As the frequency was increased, the readings were more stable, but the change in dissipation factor and capacitance with per-cent moisture was less. The results obtained at 1 kc where the change was rather great, and, at the same time, stability was achieved, are plotted in Figs. 4 and 5. A more thorough investigation of the relationship between dissipation factor and capacitance and changing per-cent saturation would probably show that, with some modification, the instrument could be adopted for this use. However, because of the simplicity, economy, and general portability of available instruments, the resistance measurement was adopted for further investigation.

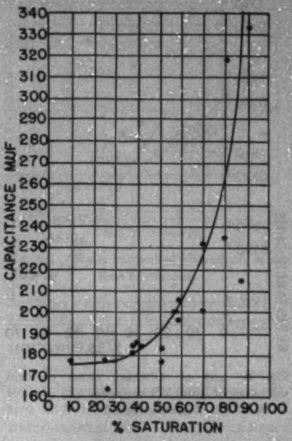


Fig. 5. Capacitance at 1 kc versus per-cent saturation of water in transformer oil, 0.003-inch fabric

Table I. Comparison of Results at Various Temperatures

Oil Temperature, C	100-Per-Cent Saturation, Parts Per Million*	Actual Water Content, Parts Per Million	Per-Cent Saturation	Resistance, Ohms
29	63	27	43	1.2×10^{10}
36	89	27	30	7.6×10^{10}
40	104	27	26	1.3×10^{11}
46	132	27	20	1.5×10^{11}
55	182	27	15	1.8×10^{11}
42	113	39	35	3.2×10^{10}
42	113	53	47	6.8×10^{9}
42	113	64	57	2.1×10^{9}

* See reference 10.

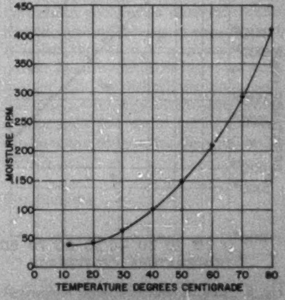


Fig. 6. Moisture saturation values of transformer oil as a function of temperature

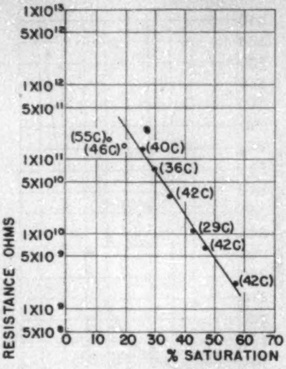


Fig. 7. Log of resistance versus per-cent saturation of water in transformer oil, effect of heating the oil

Resistance as a Function of Per-Cent Saturation

In the initial stages of the laboratory evaluation of the moisture detector, the oil temperature varied from about 27 to 31 degrees centigrade (C) during the tests. This seemed to have no effect on the resistance-versus-moisture relationship. The reason for this is directly attributable to the use of per-cent saturation rather than absolute water content. It was found that, when the temperature was increased at a constant water content, the resistance also increased. If the results were plotted in terms of absolute water content, a different curve would be obtained for each temperature level. However, the inclusion of the water content and temperature in the term per-cent saturation resulted in a single line, as shown in Fig. 3.

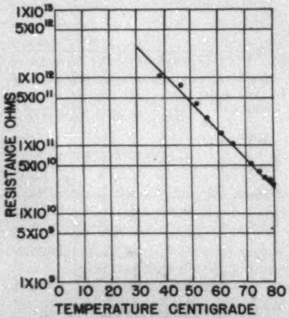


Fig. 8. Log of resistance versus temperature obtained with vacuum-dried detector in the absence of oil

To examine this effect more fully, a brief study of the influence of temperature on the results was carried out. The data obtained from this investigation are shown in Table I. The values in column two of the table were taken from a previously determined graph, reproduced in Fig. 6. The first five lines of data were obtained by heating the oil and measuring resistance at various temperatures. Reference to the last column shows that raising the temperature from 29 to 55 C increased the resistance more than a decade. If these data were plotted in terms of absolute water content, a vertical straight line would be obtained having no correlation with data obtained at room temperature. In contrast to this, the results plotted in terms of per-cent saturation gave a straight line practically identical with the one obtained at room temperature. Such a plot is shown in Fig. 7. Also included in this figure are the data on the last three lines of the table. These results were obtained by holding the temperature at 42 C and changing the water content. A comparison of Fig. 7 with Fig. 3 shows the lines are practically identical.

Some deviation toward lower resistance seems to be occurring at temperatures above 46 C. From the limited data available, this appears to be a real deviation and not merely experimental error. The results indicate that the use of per-cent saturation makes the relationship independent of temperature changes from 29 to about 46 C.

Deviations at Higher Temperatures

A clearer picture of the effect of higher temperatures on the operation of the detector can be obtained by a consideration of the consequences of heating the oil. Since the system is composed of paperlike fabric, insulating oil, and moisture, each component must be considered independent of the others and then it must be determined how each contributes to the over-all phenomenon. Since the amount of water that oil can hold in solution increases at elevated temperatures, the per-cent saturation at a given absolute water content decreases. If no other change in the system occurred while heating the oil, the resistance at any temperature would vary with per-cent saturation, as shown in Fig. 3. However, since the measured resistance is the result of a combination of several parallel resistances, its value is controlled by that component which has the lowest resistance. One of the parallel components is the fabric itself. To determine how its resistance changed with temperature, the detector was placed in an airtight chamber with no oil present and was vacuum-dried overnight. The change in resistance across the electrodes with temperature was then determined. The results are shown in Fig. 8. The limitations of time prevented a similar experiment using only dry oil, which is another parallel resistance component, but a similar decrease in resistance with temperature would be expected.

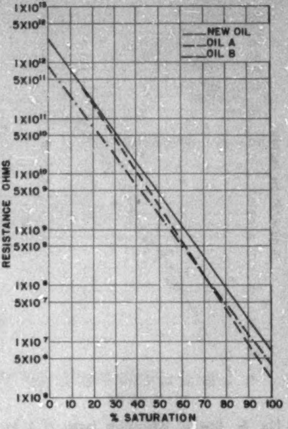


Fig. 9. Log of resistance versus per-cent saturation of water in transformer oil, comparison of new and used oils

From the experimental data available, some explanation of the deviation of results obtained at elevated temperatures from those obtained at room temperature can be presented. For example, at room temperature and 15-per-cent saturation of water in oil, the resistance was 4.1×10^{11} ohms; see Fig. 3. At the same relative water content, but at 55 C, the resistance was 1.9×10^{11} ohms; see Fig. 7. From the data plotted in Fig. 8, the resistance of the

Table II. Properties of New and Used Oils

Designation	Interfacial Tension, Dynes per Centimeter, 25 C	Acid, Milligrams KOH per Gram of Oil	100 Per-Cent Saturation,* Parts per Million, 25 C
New oil	40 to 45	< 0.01	54
Oil *A*	32	0.018	58
Oil *B*	14	0.70	126

*These values were obtained experimentally.

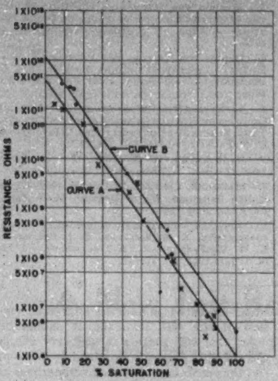


Fig. 10. Log of resistance versus per-cent saturation of water in transformer oil

Curve A—Outlet from filter press
Curve B—Equilibrium curve

dry fabric at 55 C was 2.7×10^{11} ohms. In other words, regardless of how dry the system is, its resistance cannot exceed the limitation imposed by the relationship shown in Fig. 8 for the dry paper. At higher relative moisture contents, the presence of the moisture reduces the resistance sufficiently so that the relative moisture content becomes the factor primarily controlling the resistance. Further experiments are planned to provide a more rigorous explanation of the effect of temperature on the results obtained from the detector.

OIL FROM OPERATING TRANSFORMERS

Because of the desire to obtain fundamental data on the moisture detector, most of the investigation was carried out using new transformer oil. However, since used oil will be encountered in many practical applications of the detector, it was deemed essential that some results be obtained using oil from actual operating transformers. Samples of two such oils were evaluated. The results obtained are shown in Fig. 9. For convenience, the results obtained for new oil under the same conditions are also shown. The properties of the oils are shown in Table II. An inspection of the table shows that oil *B* has a very low interfacial tension and a high acid number. It was felt that this oil represented about as poor an oil as one might encounter in the field. However, even with this oil, a constant absolute deviation of only 8 per-cent saturation occurred.

The major difficulty encountered with used oil was the necessity of determining the maximum amount of water the oil could hold at a given temperature. This quantity increases as the oil deteriorates. Experimentation on methods for estimating quickly this value for all oils has been undertaken and, so far, moderate success has been achieved. Some correlation was obtained between interfacial tension and water saturation. This work is in preliminary stages and is being actively pursued by the authors. Once a method is developed it is hoped that, from a single simple measurement on the oil, its saturation water content can be predicted. From this value, a good idea of the correction to be applied to the curve, if any, can be obtained.

FILTER-PRESS APPLICATION

To determine the characteristics of the moisture detector under conditions similar to a factory application and as part of a general program designed to devise a suitable method for the evaluation of filter papers, the plastic laboratory model of the detector was installed on a small-scale filter press. The press was a *1/200* scale model of a 15-plate 7-inch filter press. Moisture-saturated oil of known water content was filtered through the press at a flow rate of 3.6 gallons per hour which was 1/200th of the flow rate obtained with the full size press. Sampling techniques and water content measurements were identical with those previously described. The resistance-moisture relationship obtained for the outlet oil from the filter press is shown in Fig. 10, curve *A*. The equilibrium relation for the plastic detector is shown in curve *B* of Fig. 10. It should be remembered that in obtaining the latter curve, it sometimes was necessary to allow 2 gallons of oil of the same water content to flow through the gauge before equilibrium was reached. In the filtration experiment, the water content of the outlet oil changed as much as 20 parts per million during the passage of 2 gallons of oil. This was caused by the very small scale operation and the fact that the incoming oil contained as much water as it could hold in solution at the temperature of the experiment. Under actual filtering conditions, many gallons of oil of the same water content would pass through the detector. The rapid change in water content in the small-scale operation caused a slight lag in resistance readings for a given per-cent saturation, as can be seen in Fig. 10. However, once the curve was determined for a particular flow rate, it was reproducible.

After obtaining the resistance-moisture relationship, experiments were carried out to determine the rate of change of water content with time. Resistance readings were taken at frequent intervals during a filtering operation and from them the per-cent saturation of the oil was determined. The water content of the oil was calculated and plotted against time. During the test three water contents were determined by sampling the oil and titrating with Karl Fischer reagent. Fig. 11 shows a graph of the water content versus time. The circles are for values determined by the moisture detector and the squares for values obtained by titration. The close agreement of the data obtained from the detector with that obtained by chemical means is evident. Thus, by simply measuring resistance at frequent intervals, the water content of the outlet oil was monitored very closely. If the data shown in the curve were obtained by sampling and titrating, ten samples would be needed. Several different filter papers were evaluated in the foregoing manner to determine their relative ability to dry oil efficiently. From the results obtained, a clear differentiation among the filter papers was obtained. It is felt that on a large press where a greater quantity of oil of identical water content will pass through the instrument, the closeness of approach to equilibrium will be such that the required accuracy can be obtained by using the equilibrium curve.

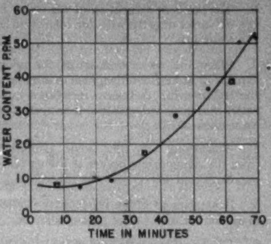


Fig. 11. Water content of filtered oil versus time of filtration

□—Water content obtained by titration
○—Water content calculated from resistance reading of detector

Conclusions

The data presented in this paper have led to some important conclusions:

1. The resistance, as measured with the moisture detector described in this paper, can be sensibly correlated with the relative water content of the oil passing through it.

2. The per-cent saturation of the water in the oil varies directly with the logarithm of the resistance.

3. When the water content of the oil is expressed in terms of per-cent saturation, the relationship between resistance and water content is independent of temperature up to about 46 C.

4. Indications are that some deviation from the straight line occurs at higher temperatures and lower per-cent saturation.

5. It appears as if the instrument is adequate for use with rather poor-quality oil.

6. The detector can be used to measure the water content of oils continuously in filtering operations and other applications involving flowing oil.

References

1. A New Method for the Analytical Determination of the Water Content of Liquids and Solids, **Karl Z. Fischer.** *Angewandte Chemie*, Berlin, Germany, vol. 48, 1935, pp. 394–96.

2. Fischer Volumetric Determination of Water. **E. G. Almy, W. B. Griffin, C. S. Wilcox.** *Industrial and Engineering Chemistry*, Easton, Pa., Analytical Edition, vol. 12, 1940, pp. 392–96.

3. **G. R. Cornish.** *Plastics*, New York, N.Y.,1946, pp. 99–103.

4. Improved Test Cell Stabilizes Oil Measurements, **E. C. Schurch.** *Electric Light and Power*, Chicago, Ill., vol. 30, July 1952, pp. 98–100.

5. Method and Mechanism for Determining the Moisture Content of Paper or Other Material, **A. Allen.** *U. S. Patent 1,708,079*, May 1, 1925.

6. Device for Testing the Insulating Properties of Insulating Liquids Such as Oils, **C. Dantsizen.** *U. S. Patent 1,721,374*, March 29, 1926.

7. Apparatus for Determining the Moisture Content of Wood or Other Materials Through Their Electrical Conductivities, **C. B. Limbrick.** *U. S. Patent 1,890,545*, Aug. 23, 1930.

8. Hygrometer, **H. S. Polin.** *U. S. Patent 2,015,125*, Jan. 11, 1934.

9. Moisture Content Recorder for Gases Under Pressure, **W. M. Deaton.** *U. S. Patent 2,629,253*, June 1951.

10. Gas and Moisture Equilibriums in Transformer Oil, **R. B. Kaufman, E. J. Shimanski, K. W. MacFadyen.** *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 312–18.

◆

Discussion

F. C. Doble (Doble Engineering Company, Belmont, Mass.): This paper presents some very interesting laboratory results on measuring moisture contents in a flowing stream of oil. Considerable experience with the measurement of water content in flowing streams of oil, both in the laboratory and in the field, indicate that there are many important factors that have not been mentioned in this paper but which nevertheless must be taken into consideration. For example, the rate of flow, size of electrodes, thickness of the fabric, pressure on the fabric, and the effect of insulation in parallel with the electrodes must be considered, particularly when the d-c insulation resistance measurements are used to evaluate the moisture contents.

Contrary to the authors' statements we have found that 60-cycle power factor gives the most reproducible and repeatable results in the field because the cell and instrument can be so arranged as to be relatively unaffected by the factors mentioned in the foregoing. In addition, d-c measurement on the order of 100,000 megohms, such as are required for low saturation values, are exceedingly difficult to make accurately under field conditions.

While it might be of general interest to know the moisture-sensitive fabric used by the authors, we have found that many different materials can be used but the sensitivity and calibration may depend upon the material chosen.

G. M. L. Sommerman (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The electrical test method for determining the moisture content of dielectric liquids described in this paper is based on the Schurch method (reference 4 of the paper). The authors have added certain refinements such as ensuring that equilibrium is attained between the moisture in the liquid being tested and the moisture in the moisture-sensitive fibrous detecting element. Also they seem to favor the use of d-c measurements on the detecting element instead of dissipation factor measurements at 1,000 cycles per second (Schurch measured dissipation factor at 60 cycles per second). It is true that the electrical properties of many materials are more affected by moisture as the measuring frequency is lowered, and are affected most at zero frequency (d-c). However, sources of error in the measurements must be considered also. Some of these are mentioned by Schurch.

Another source of error in measuring the d-c properties of the moisture-sensitive element (paper-like fabric) is that caused by the resistance of the dry fabric in parallel with the resistance due to the moisture picked up by the fabric from the liquid. The authors discuss this to some extent in the section entitled "Deviations at Higher Temperatures." They correctly refer to their data on dry-element resistance as a function of temperature, Fig. 8, and show that this explains the low points on the resistance-versus-moisture calibration curve of the element, Fig. 7, at temperatures of 46 C and above. Further analysis of these data shows that conclusion 3 is not complete when it states that ". . .the relationship between resistance and water content is independent of temperature up to about 46 C." Following through the authors' analysis in the section just mentioned, we can compute, as a function of element temperature, the element resistance from the parallel resistance of the element with moisture present, Fig. 7, and the dry-element resistance, Fig. 8. This leads to the following: 1. at zero moisture content in Fig. 7, the data in Fig. 8 indicate the intercepts of the corrected calibration curves for various element temperatures; 2. some approximate intercept values for these corrected curves in Fig. 7 are 30×11^{11} ohms at 30 C, 10×10^{11} ohms at 40 C, 5×10^{11} ohms at 50 C, and 1.5×10^{11} ohms at 60 C; 3. from these intercepts, a family of corrected calibration curves may be computed and drawn to merge with the present single curve in Fig. 7 at intermediate values of moisture contents. At successively higher temperatures, the successive curves have lower intercepts (as indicated in the foregoing) and merge with the original curve at successively higher percentages of water. Thus the true conclusion is that the curve in Fig. 7 holds at 30 C for all water contents, holds at 40 C for all water contents above 12 per cent, holds at 50 C above 24 per cent, holds at 60 C above 36 per cent, etc. (percentages are approximate per-cent saturations of water in oil).

It is important to note that the amount of water in oil at 100-per-cent saturation depends on the degradation products in the oil, Table II, as well as the oil temperature, Table I. It is encouraging to note that the authors believe that there may be a correlation between the amount of water in a water-saturated oil and the interfacial tension of the oil. However, it should be noted that still another oil property affects the amount of water it can hold in solution, i.e., the percentage of cyclic hydrocarbons in oils. This is shown by the data in Fig. 6 of reference 10.

M. F. Beavers, E. J. Shimanski, and **E. F. Timpane:** Each of the factors mentioned as important by Mr. Doble have been taken into consideration by us, and in some cases have been mentioned in the paper. Rate of flow was dealt with in the section entitled "Filter-Press Application" because this was the only instance in which flow affected the results. It should be mentioned that the effect of flow rate in this case was due to a nonequilibrium condition in the moisture-sensitive fabric and it would be encountered regardless of the measuring circuit. While it is quite true that the size of electrodes will change the resistance, this is certainly no disadvantage since the electrode size is a constant value.

The effect of electrode spacing on the results is discussed in the section entitled "Fundamental Studies." We have found very little effect resulting from pressure variations. We have obtained essentially the same results from a model which was bolted together, and from the glass model in which the tension was accomplished by a small spring. As for insulation in parallel, it is of course true that the parallel resistance must be greater than that which is being used to detect moisture. In our case the use of teflon as a spacing material was dictated by its extremely high insulation resistance over the entire range of per-cent saturation. This was pointed out in the section entitled "Description of the Moisture Detector."

Our inability to use 60-cycle power-factor measurements was probably due to the particular dimensions and construction of the detector as well as the equipment available for measuring this quantity. No extensive investigation was made using power-factor measurements. Since the literature contains no instance of a thorough study conducted on an instrument such as is described in the paper, the decision to use resistance measurements was based solely on our own experimental data as determined in the laboratory.

With regard to instrumentation, we feel that the resistance measurement, which can be carried out with a battery source up to

Vol
Pt.
19

100,000 megohms, is about as simple to measure in the field as any electric measurement which will reflect changes in water content.

Regarding Dr. Sommerman's comments, we do not feel that our method is based on the method described by Schurch any more than it is based on the methods given in references 5 through 9 of the paper. Schurch seems to be dealing with a nonhomogeneous system of water and oil which he states contains some free water. Free water causes an immediate and very great drop in resistance on our detector which results in a final resistance too low to be measured on our bridge. If oil such as Schurch described were to be passed through our detector, the result would be impossible to interpret in terms of parts per million water such as is done in his article. The use of perforated plates with paper-like material between them is similar in both instruments. This, however, is the only apparent likeness.

Dr. Sommerman's discussion of conclusion 3 of the paper is quite correct, and we agree that it should have been modified in the manner referred to by Dr. Sommerman. It should, however, be pointed out that the deviation of results at higher temperature is not necessarily a disadvantage resulting from the use of a d-c measurement, as Dr. Sommerman implies, but might also occur using dissipation factor.

The relationship between interfacial tension and water saturation in oil seems to be valid over a limited range and it is hoped that this range will prove sufficient for most applications. As Dr. Sommerman mentioned, it is well to keep in mind when discussing water in transformer oil that structural differences will result in different 100-per-cent saturation levels.

A New Test Board for Portable Watt-Hour Meter Standards

F. J. LEVITSKY
ASSOCIATE MEMBER AIEE

THE Central Laboratory of the New England Power Service Company maintains the watt-hour standards for the New England Electric System. Until very recently, the methods and equipment used were similar to those found in many utility laboratories. Primary standards were standard cells, which, through potentiometers, volt boxes and shunts, established the calibration of laboratory standard wattmeters. The wattmeters, in turn, served as transfer instruments from d-c reference to a-c working values. A good-grade pendulum clock served as time standard.

In common with many other laboratories operated by both utility companies and manufacturers, it has been found that in spite of the fact that the basic reference for the watt-hour is the combination of dynamometer wattmeter and pendulum clock, a more convenient, secondary standard is provided by carefully selected watt-hour meters. The use of an indicating wattmeter relies on the human element to integrate and thus arrive at watt-hours. The pendulum clock is considered a satisfactory source for the time element in such integration. Nevertheless, the tester must be able to hold a constant watt value through the predetermined time interval. Automatic watt regulators have made this task easier, but such devices correct rather than compensate for fluctuations. The use of a watt regulator with a stable voltage supply, however, may reduce all deviations to negligible magnitudes.

Investigations in utility laboratories and by manufacturers over many years have shown that calibrations made with indicating instruments show a greater variability than those obtained through the use of watt-hour meter standards.[1] Recent work by the National Bureau of Standards also corroborates this conclusion.[2]

The use of watt-hour meter standards as references for the calibrations of field watt-hour meter standards simplified to a considerable degree the task of calibration but did not improve the basic accuracy achieved. The improvement in watt-hour meter characteristics during the past few years, however, made such an improvement a practical necessity. Difficulties in calibrating the reference standards existed, and as an attempt was made to narrow the tolerances in the watt-hour values, these difficulties increased. Consideration of the variability experienced in laboratory calibrations led to an investigation of the following three factors: 1. The adequacy of the time element as derived from a pendulum clock. 2. The consistency of the indicating wattmeter calibrations. 3. The dependability of the test procedure used to calibrate the reference standards.

It was concluded that the pendulum clock was an adequate standard for time measurement. Two methods were available to obtain the desired 1-minute interval: contacts operated by the mechanical escapement in the clock and a photoelectric pickup actuated by the pendulum. Thus a cross check on the accuracy of contact mechanisms is always available.

Indicating wattmeter calibrations against the potentiometers gave very consistent results. Provisions have been made for wattmeter calibrations without moving the instrument so that in case of doubtful test results reference can immediately be made to primary standards.

The reference watt-hour meter standards showed minor, but irritating, deviations among themselves particularly in calibrations at lagging power factor (pf). Wattmeter tests of these standards also showed similar errors which were minor in so far as commercial standards were concerned, but annoying from the laboratory point of view. To minimize these errors, test procedure and equipment were reviewed. The aim was to stabilize all conditions which might influence the reference standards, to eliminate as far as possible the errors due to the human element, and thus reduce inaccuracies of the standards to those inherent in selected watt-hour meters operating under ideal conditions. Then, by establishing a procedure for testing both reference standards and field standards that would be simple and trouble-free it was believed that a level of accuracy could be reached that would satisfy present-day requirements.

Using the best ideas which could be found among manufacturers and utilities a new test board was designed and constructed. This board employs selected watt-hour meter standards operating under as near ideal conditions as possible, and provides a simple, accurate means for

---

Paper 55-150, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 19, 1954; made available for printing November 29, 1954.

F. J. Levitsky is with the New England Power Service Company, Worcester, Mass.

Invaluable aid in the design and construction of the decade counter and associated photoelectric pickups was obtained from G. N. Burkhart of the Duncan Electric Manufacturing Company. G. A. Palmer and L. G. Eaton of the New England Power Service Company aided in the over-all design of the test board.

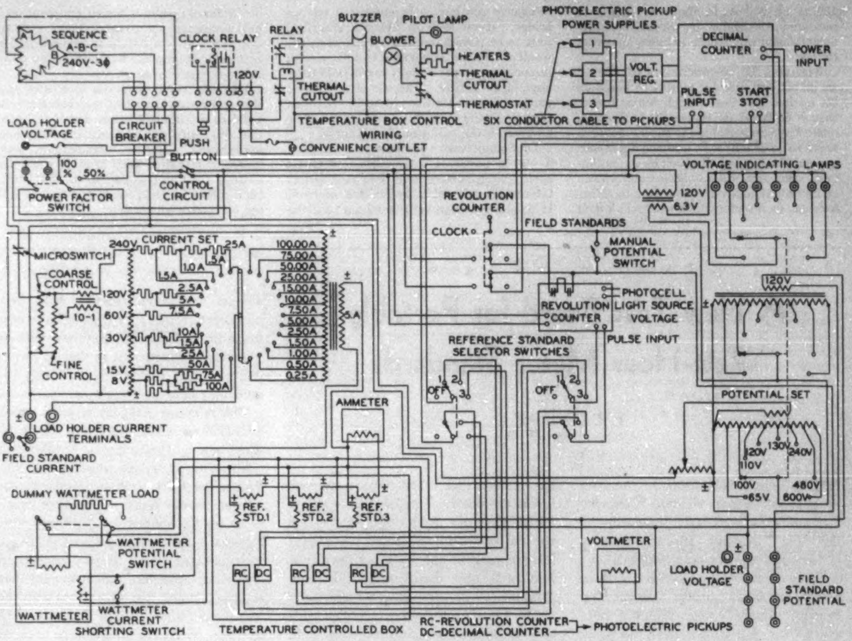

SEQUENCE
A-B-C
240V-3Φ
A
B
C
CLOCK RELAY
RELAY
BUZZER
PILOT LAMP
BLOWER
HEATERS
THERMAL CUTOUT
THERMOSTAT
THERMAL CUTOUT
120V
PHOTOELECTRIC PICKUP POWER SUPPLIES
1
2
3
VOLT. REG.
DECIMAL COUNTER
POWER INPUT
PULSE INPUT
START STOP
SIX CONDUCTOR CABLE TO PICKUPS
TEMPERATURE BOX CONTROL WIRING
CONVENIENCE OUTLET
LOAD HOLDER VOLTAGE
CIRCUIT BREAKER
PUSH BUTTON
100 %
50%
POWER FACTOR SWITCH
CONTROL CIRCUIT
VOLTAGE INDICATING LAMPS
120V
6.3V
REVOLUTION COUNTER
CLOCK
FIELD STANDARDS
MANUAL POTENTIAL SWITCH
MICROSWITCH
240V
CURRENT SET
25A
.5A
1.0A
1.5A
2.5A
5A
7.5A
10A
15A
25A
50A
75A
100A
COARSE CONTROL
10-1
120V
60V
30V
15V
8V
FINE CONTROL
100.00A
75.00A
50.00A
25.00A
15.00A
10.00A
7.50A
5.00A
2.50A
1.50A
1.00A
0.50A
0.25A
5.A
PHOTOCELL LIGHT SOURCE VOLTAGE
REVOLUTION COUNTER
PULSE INPUT
120V
REFERENCE STANDARD SELECTOR SWITCHES
OFF 1 2 3
OFF 1 2 3
AMMETER
POTENTIAL SET
120V
130V
240V
110V
100V
65V
480V
600V
LOAD HOLDER CURRENT TERMINALS
FIELD STANDARD CURRENT
DUMMY WATTMETER LOAD
REF. STD.1
REF. STD.2
REF. STD.3
WATTMETER POTENTIAL SWITCH
VOLTMETER
RC DC
RC DC
RC DC
LOAD HOLDER VOLTAGE
FIELD STANDARD POTENTIAL
WATTMETER
WATTMETER CURRENT SHORTING SWITCH
TEMPERATURE CONTROLLED BOX
RC-REVOLUTION COUNTER
DC-DECIMAL COUNTER
PHOTOELECTRIC PICKUPS

Fig. 1. Complete schematic circuit diagram of test board

calibrating the reference standards as well as the standards brought in from the field.

Circuits and Equipment

The reference meters used on the new board to test field standards are three special watt-hour meter standards which operate continuously and are housed in a temperature-controlled box. The customary photoelectric devices on the reference standards are used to test field standards. In addition, each reference standard has a second photoelectric system which counts fractional as well as whole revolutions to permit their test by indicating wattmeter. Through the use of precision transformers, calibration of the reference standards at only two points suffices, one at unity pf and one at 50 per cent (%) lagging pf. A complete schematic wiring diagram of the test board is given in Fig. 1, A general view of the front of the board is shown in Fig. 2.

By employing three reference watt-hour meter standards a quick intercomparison of the standards can be made. It is assumed that as long as the three standards agree there is little likelihood of significant errors in calibration. It is probable that if only minor deviations occur among the three standards their average will not significantly deviate from true watt-hours. However, any such deviation would call for a wattmeter check.

To eliminate ambient temperature errors, the three reference standards are installed in a temperature-controlled box and maintained at a temperature of 100 degrees Fahrenheit. This temperature-controlled box consists of a box within a box, with the reference standards in the inner compartment, as shown in Fig. 3. Heaters are mounted at the bottom between the boxes, while a blower and baffle serve to circulate the air. Safety devices guard against faulty thermostat operation or failure of blower. Tests show that temperature in the inner compartment is maintained well within 2 degrees Fahrenheit.

Fig. 2. General view of test board showing control panel, temperature-controlled box mounted under table, and indicating wattmeter with its associated test board at left

Vol. 7
Pt. 1
1955

Fig. 3. View of interior of temperature-controlled box showing reference standards and associated photoelectric pickups

Fig. 4. Enlarged section of a reference standard disk showing teeth cut in edge

To remove self-heating errors, the references standards are run continuously. Vibration errors are minimized by shock mounting of the temperature-controlled box. Friction errors have been lessened by removal of registers in the reference standards, and intercoil errors are absent since each standard has a single current coil and a single potential coil. The use of watt-hour meter standards as a reference for testing field standards presented no problem, since the normal test equipment could be used. This consists of a photoelectric tester which permits the standards under test to run during the time required for a predetermined number of whole revolutions of the reference standard. The register reading of the standard under test indicates its performance.

With the elimination of the register, a major problem was the calibration of the reference standards. It was possible to use the photoelectric pickups installed in the standards to obtain a count of full revolutions only. A current method uses this count of full revolutions in conjunction with a chronograph, giving precise time measurement, to compute watt-hours with time as the variable. Theoretically, this method gives good results; but the use of the chronograph is tedious and there is a reluctance on the part of the tester to make calibrations by this method.

In spite of the accuracy of the chronograph it was believed that a definite time interval could be fixed more precisely than an indeterminate time interval could be measured, and did not introduce a human error in such time measurement. Thus it was decided, for calibration of the reference standards, to retain the 1-minute interval and to measure both partial and whole revolutions of the disks by employing an additional photoelectric pickup in each standard and an electronic counter.

To actuate the photoelectric pickup, 200 evenly spaced teeth approximately 1/16 inch deep were cut in the edge of the disks in each reference standard. A section of the edge of one disk is shown in Fig. 4. Since the aperture between adjacent teeth would not transmit sufficient light to actuate the phototube, a metal plate in which a 1/2 by 1/16-inch lattice had been cut was mounted over a segment of the meter disk. The lattice openings match the slits in the disk to form a shutter. The photoelectric system has the light source above the shutter and the phototube below. As the disk rotates, the light intensity on the phototube varies from a small value when the teeth in the disk are directly under the slots in the plate to a maximum value when the slots in the disk are in line with the slots in the plate, thereby giving 200 output pulses per revolution. The magnitudes of these pulses are still too small to be used directly, so a 2-stage amplifier is mounted on the pickup chassis. The amplified pulses are then fed to a high-speed electronic counter. This counter consists of wave shaper, amplifier, and four decade-counting units which give a direct reading of the number of pulses. The rated maximum input of the decade counting units is 40,000 pulses per second, while the maximum input rate in this circuit is 55 pulses per second.

To calibrate the references standards, which run continuously, the power is held constant and the clock relay contacts operate the start-and-stop circuit of the decade counter. A total reading on this counter will give, when divided by 200, whole revolutions and partial revolutions of the reference standards to the nearest 0.005 revolution.

The accuracy of the decade-counting units is absolute ±1 count. Therefore, percentage errors can be made small by making total number of counts large. Errors in starting and stopping have been reduced to a minimum by the use of a start-stop relay with a rated operating time of $^1/_2$ millisecond. Since the minimum time between incoming pulses is 18 milliseconds, the possibility of obtaining spurious counts by input relay operation is small.

Current and voltage for the field standards are obtained with Uniload equipment. Current from 0.25 ampere to 100 amperes and potential from 65 to 600

Table I. National Bureau of Standards Test Data on Precision Transformers

Current Transformer, Burden: 4.07 Volt-Amperes, 0.68 PF, Lagging Current				Potential Transformer, Burden: 51.9 Volt-Amperes, 0.259 PF, Lagging Current			
Primary Amperes	Secondary Amperes	Ratio Correction Factor	Phase Angle, Minutes	Primary Volts	Secondary Volts	Ratio Correction Factor	Phase Angle, Minutes
0.25	5.00	0.9998	+0.3	65	120	1.0002	0
0.50	5.00	0.9998	+0.5	100	120	1.0001	+1
1.00	5.00	0.9998	+0.6	110	120	1.0000	+1
1.50	5.00	0.9999	+0.7	120	120	1.0000	0
2.50	5.00	0.9999	+0.4	130	120	1.0001	0
5.00	5.00	1.0000	+0.4	240	120	0.9999	0
7.50	5.00	1.0000	+0.5	480	120	0.9999	0
10.00	5.00	0.9999	+0.4	600	120	0.9997	0
15.00	5.00	1.0000	+0.4				
25.00	5.00	1.0001	+0.4				
50.00	5.00	1.0001	+0.4				
75.00	5.00	1.0000	+0.1				
100.00	5.00	1.0001	+0.2				

volts are available, as shown in Fig. 1. Selector switches not only control the desired current and potential but at the same time connect proper primary taps on the precision current and potential transformers. The reference standards are supplied from the secondaries of these transformers and, hence, operate with constant magnitude voltage and current.

The precision transformers are compensated for the secondary burden actually used. Tests on these transformers at the National Bureau of Standards determined ratio and phase-angle errors with rated burden as shown in Table I. Calculation of combined transformer errors at the commonly used calibration points shows a maximum of 0.03% at unity pf and 0.04% at 50% lagging power factor. Considering the performance repeatability of modern portable watt-hour meter standards, these errors are considered insignificant and no transformer corrections are applied in routine calibration.

Use of this new equipment has produced more consistent results than were believed possible in the calibration of watt-hour standards. A load holder and high-capacity power supply to the test board have, no doubt, assisted in this achievement.

Operation

As outlined in the foregoing, absolute calibration of the reference standards is accomplished by holding the wattage constant and measuring revolutions and partial revolutions for a fixed time interval with the decade counter. Since the reference standards always operate with the same magnitude voltage and current, it is necessary only to calibrate them at two points, unity pf and 50% lagging pf. Such tests are made weekly.

Even with constant load the speed of the meter disk varies during each revolution, therefore variations can occur in calibrations requiring measurement of partial revolutions. Because of this and the possibility of a ±1 count error in the decade counter, no single test result on the reference standards is used. The posted corrections for these standards are the average of a large number of tests.

Provisions have been made to obtain relative accuracy checks on the reference standards as well as a check on counter accuracy. Switching is provided, as shown in Fig. 1, to allow the contacts of the revolution counter to operate the start-stop circuit of the decade counter. With the switch in this position, a check on the counters can be made by operating each counter from the same reference standard. For this condition, a 1-revolution setting on the revolution counter should produce a count of 200 ± 1 count on the decimal counter, regardless of the speed or condition of the reference standard used. Similarly, a 10-revolution setting should produce a count of 2,000 ±1, giving a check on both counters.

If the decade counter is connected to one reference standard and the revolution counter to another, the total count shown on the decade counter gives the relative accuracy between the two standards. Comparing the indicated difference in performance between the standards with the difference between the posted corrections provides a quick check on the relative accuracy of the standards. These checks are made daily.

Field standards are calibrated in the normal manner with the use of a standard photoelectric tester. All such tests use only full revolutions of the reference standards, so errors due to variation in speed of the reference standard disks during any one revolution are eliminated.

The stability of the reference standards is considered excellent. Calibration data posted to the nearest 0.05% show no change in the average calibration of the reference standards at either unity or lagging pf during the first 3 months of use. During this period 52 tests were made on each reference standard at both pf values.

Conclusion

With the described equipment, a very high degree of consistency among reference watt-hour meter standards has apparently been achieved. Calibrations of these standards indicate a stability considerably better than had been indicated by previous experience. In many ways the described equipment is similar to that used by other laboratories. The contribution here is believed to be the development of what is considered a new technique for calibrating watt-hour meter standards by employing two photoelectric systems. The use of these two systems makes it possible to calibrate any one reference standard by laboratory indicating instruments, to intercompare any two of the three reference standards, to cross-check accuracy of the mechanical and electronic counters, and to calibrate with the required degree of accuracy standards brought into the laboratory.

References

1. A Precision Rotating Standard for the Measurement of Kilowatt-Hours, **J. H. Goss, A. Hansen, Jr.** *AIEE Transactions* (*Electrical Engineering*), vol. 59, July 1940, pp. 412–16.

2. Improved Method for Precise Rate-Measurement of Watthour Meters. *Technical News Bulletin*, National Bureau of Standards, Washington, D. C., vol. 38, Sept. 1954, pp. 132–33.

◆

Discussion

J. H. Spencer (General Electric Company, Somersworth, N. H.): The value of maintaining three precision watt-hour meter standards for calibrating portable watt-hour standards has been recognized for a number of years. Precision watt-hour standards complete with a constant-temperature box are commercially available and are being used by numerous electric utility companies.

It is customary to standardize the precision watt-hour meter standard by energizing the standard with a constant wattage and then measuring the time required for a definite number of full revolutions of the disk of the standard. A laboratory-type indicating wattmeter is used in conjunction with manual or automatic control to maintain a constant wattage. To obtain maximum accuracy, the wattmeter calibration is normally checked before and after each use. This check is made with direct current using potentiometers and standard cells.

Chronographs are used to measure the time required for a definite number of full disk revolutions. Briefly a chronograph for this purpose consists of a paper tape driven at constant speed and two recording pens. One pen is used to record time signals from an accurate timing source such as a pendulum clock. The trace from this pen appears as a straight line with "pips" spaced 1 second, or some other time interval, apart. The second pen is actuated by a photoelectric system operating through a hole in the disk of the standard being checked. By a comparison of the two traces an accurate determination of the time for one or more full disk revolutions can be made. By using a 4-pen chronograph, a simultaneous check of all three precision watt-hour meter standards can be made.

Although accurate measurements can be made with a chronograph, Mr. Levitsky has recognized that its use is somewhat tedious and subject to human error. He has developed a novel system in which the fractional and whole revolutions of the disk are determined electronically for a definite time interval. Since the disk of the precision watt-hour meter standard does not rotate at a uniform speed during any one revolution, an error will be present in any calibration that includes partial revolutions. The magnitude of this error will, of course, decrease as the number of full revolutions checked is increased. With a sufficiently large number of revolutions the error would be negligible.

An alternate means for accomplishing Mr.

Levitsky's objectives has been tried and has proved highly accurate. In this system, like the chronograph, the time required for a definite number of full revolutions of the disk is measured. Instead of using the photoelectric signal timed to disk revolutions to actuate a chronograph pen, it is used to start and stop an electronic time interval meter. An intermediate device is used ahead of the time interval meter to start the time interval meter and to stop it after a predetermined number of revolutions of the disk. The time interval meter that has been used reads directly in milliseconds and has a time base stability of 5 parts per million.

G. F. Austin (Public Service Company of New Hampshire, Manchester, N. H.): The comparatively new methods and equipment as described in Mr. Kevitsky's paper on the standardization of the watt-hour are undoubtedly a tremendous improvement over the older methods. There is no doubt but that the use of a rotating watt-hour standard under ideal conditions will give excellent repeatability.

I have recently had the opportunity of inspecting the seeing in operation the equipment described in this paper. The design and construction of the New England Power Service Company's new standard test board surely represents the most modern thinking in this field.

Our company uses the same basic principle and similar equipment. Probably the major difference between their equipment and ours is the use of the chronograph to measure the calibration of the primary rotating standards. Mr. Levitsky mentions this technique and states that theoretically this method is probably acceptable. After 3 years' experience with this method, we are able to report that in actual practice it has proved satisfactory. However, we agree that this method is time-consuming and tedious. If, in our opinion, the electrical counter method proves it will give us the same degree of accuracy as we now experience, we will undoubtedly convert to it some time in the future. The controversy over the two methods centers around the age-old argument that the materials in the disk of a watt-hour standard will probably not be uniform, and this together with the anticreep holes will tend to cause the disk not to rotate at a uniform speed throughout a single revolution. Therefore, to secure the best results, a calibration should never be run on a standard without using a complete number of revolutions of the disk. The chronograph method which we use employs the use of complete standard disk revolutions. The counter method which is described in this paper uses a number of revolutions and a fraction of a revolution. However, again I believe that the New England Power Service Company has proved to their satisfaction that in actual practice this is not a consideration. At frequent intervals, we cross-check calibrations with the New England Power Service Company using a transfer watt-hour standard. The difference in the two evaluations of the kilowatt-hour seldom disagrees over 0.1% at unity pf.

We, like Mr. Levitsky, agree that these primary rotating standards should run at ideal conditions in a temperature-controlled box free from vibration. From a theoretical consideration, we think that perhaps it is unwise to insert an air-circulating motor into a temperature-controlled box. We theorize here that the benefits derived from circulated heated air may be offset by the undesirous motor vibrations set up in the equipment.

The author has mentioned the use of automatic load holder when standardizing the watt-hour meters from the dynamometer wattmeter and standard time impulses. We have such a load holder and are reluctant to use it because of some recent experiences we have had with it. Differences of as much as 0.1% from manual checks have been experienced when using the load holder. A thorough investigation showed that using our particular type of load holder we were getting an additional 2% wave-form distortion caused by a feedback of a high-frequency harmonic generated in the load-holding equipment. This wave-form distortion is very visible on a scope, and until such time as we are able to filter out these undesirable harmonics we will continue to hold load manually.

As Mr. Levitsky has pointed out, the elimination of high resistance contacts in switches is a basic consideration in the design of such a board. It is our opinion that the substitution of switches with silver-plated straps wherever possible should be employed. The alternative is to use the best grade of switch with capacity of many times the current it will be called upon to carry. Lowering the contact resistance in this way is conducive to the best results, particularly at 50% pf checks.

My congratulations to Mr. Levitsky and his associates for what I consider a contribution to better standardization of the watt-hour within the utilities. From my experience with their company, I know they are always ready and willing to pass on their knowledge to other utilities.

W. S. Pritchett (University of California, Berkeley, Calif.): This paper presents new and effective uses of a modern electronic decimal counter in the calibration of primary reference rotating standard watt-hour meters. It makes possible the use of a fixed time interval and at the same time permits use of standards without registers. It affords an accurate means for measuring the fractional turns of the disk by photoelectric means in the absence of the register and thus eliminates friction errors caused by registers.

The author refers to the variability experienced by others in successive calibrations of the rotating standard made with indicating instruments. Mention of the range of this variability with his new equipment in the weekly tests made during the first 3 months might have been made in the paper to give a quantitative idea of the consistency of his calibrations. Also mention is made of the variation of disk speed during each revolution at constant power. Has the magnitude of this variation been measured?

In connection with the schematic diagram of the test board shown in Fig. 1 near the center of the diagram, it appears that a connection has been accidentally omitted between the upper switch arm and the center contact of the center switch of the 3-position switch assembly. Also it is noticed in the lower left corner of the figure that provision is made for a dummy wattmeter load in the potential burden circuit of the precision potential transformer, but none is provided in the current burden circuit of the precision current transformer. Instead, a short-circuiting switch is used to remove the current coil from the circuit when the laboratory standard wattmeter is not used. The wattmeter in Fig. 2 seems to be a Weston model *326* which has a current coil burden of 0.91 volt-ampere at a pf of 0.88. This is an appreciable fraction of the current circuit burden of 4.07 volt-amperes at 0.68 pf stated in Table I. Closing the switch would change the burden on the precision current transformer. This would change the ratio correction factor and possibly the phase angle slightly. Although it is noted that the transformer errors are very small and, as the author points out, these need not be considered in routine calibration of portable rotating standards, it appears that in the interests of reducing known possible sources of error to a minimum, a dummy burden for the current coil might be provided.

In connection with this burden question our laboratory has had successful experience in the use of dummy burdens on precision transformers. The a-c standardizing table in the Physical Standards Laboratory of the University of California (Berkeley) contains precision multirange transformers employing the uniload system to which the author refers in his paper and which was described some years ago by its inventor, Otto A. Knopp.[1] Acknowledgement for the final design of this standardizing table embodying the Knopp method and transformers is given to F. H. Krauss and P. T. Reimann of Electrical Facilities, Inc., Oakland, Calif. In this table the transformers have a high order of accuracy similar to that shown in Table I of the paper, and they are not readily affected by burden changes. Nevertheless, for the reason mentioned earlier, provision is made for inserting an equal dummy burden whenever a current coil or voltage circuit of an instrument or meter is removed. This maintains in use the same burden at which the transformer was certified by the National Bureau of Standards. This detail is especially desirable whenever the transformer is also used, as it is in the University table, as a standard to check the phase and ratio errors of other instrument transformers. This is evident because in such an application the demands for accuracy are even greater and during this operation all of the several instruments and meters normally in the burden circuits are removed.

Reference

1. Some Applications of Instrument Transformers, Otto A. Knopp. *Electrical Engineering*, vol. 55, no. 5, May 1936, pp. 480–89.

F. C. Holtz (Sangamo Electric Company, Springfield, Ill.): The care and good judgment which the author has used in the selection of methods and equipment for the testing and calibration of watt-hour meter standards has greatly facilitated the procedure and reduced personal and instrumental errors. This arises from the use of precision instrument transformers and the uniload system originally used and advocated by O. A. Knopp of the Pacific Gas and Electric Company.

A second important step, in my opinion, is the elimination of observational errors through the use of decade counters which enables comparisons to be made while the meters are running continuously under substantially constant load. Testing laboratories are well aware of these advantages and are rapidly changing over to the types of equipment which are advocated by the author.

While the present equipment greatly reduces personal and observational errors, it does not remove the desirability of using high-precision transfer instruments such as the electrodynamometer watter employed at the Bureau of Standards or the Shotter-Elliot a-c d-c comparator.

A high degree of consistency among watt-hour meter standards is demonstrated by the author's equipment. The achievement of such results follow from the proper design and construction of the instruments themselves, to which the manufacturers have devoted much attention in recent years. Alterations in the adjustments of portable watt-hour meters made by the laboratory at the time of test are frequently accompanied by relatively long periods of stabilization. It would be interesting to learn how such problems are handled by the author.

F. E. Davis (Anchor Manufacturing Company, Manchester, N. H.): Mr. Levitsky, while limiting his application to watt-hour meters, has presented a novel means of intercomparing, with a high degree of preciseness, the rotation of integrating meters. In watt-hour standard testing and calibration, the particular method of comparing the transfer standards with the standard wattmeter is old in principle but new in technique, as he points out. Of particular interest to those familiar with the problems involved is the auxiliary equipment utilized and the degree of precision and repeatability of results obtained with it, as well as the stability with respect to time of the modern watt-hour standards utilized as the transfer references.

Mention is made of a load holder which presumably holds the load constant at any selected value at both unity and 50% pf condition. It would be interesting to know more in detail how this is accomplished.

Thomas L. Zpaf (National Bureau of Standards, Washington, D. C.): The use of the watt-hour meters as reference standards for the testing and calibration of field standards in place of the method using a wattmeter and time signals is rapidly gaining favor with those who are concerned with improving the precision of their tests without increasing the labor and time required per test. Mr. Levitsky's paper gives an excellent method of counting partial turns of a watt-hour meter utilizing modern, high-pulse rate, electronic counters. The variation in speed of rotation of the disk of a watt-hour meter at constant load sets a limit to the precision possible in a single observation during the calibration procedure on the reference standards. Variations in the speed of the disk can be the result of a number of asymmetrical attributes of the disk. One definite source of periodical variation arises from the existence of holes in the disk, whether they be for anticreep purposes or for photoelectric counting. It may be possible to substitute disks having no holes for the ones presently used to eliminate this source of error. If this were done, the counting of an integral number of turns of the disk would necessarily have to be done by means other than the photoelectric pickoffs operating in conjunction with the holes. For the testing of other watt-hour meters the electronic counter described might be adapted to operate the start-stop relay for a predetermined number of pulses derived from the serrated edge of the disk of the standard, thereby eliminating one photoelectric system. In reality, the adaptation to the counter may be quite complex, and in view of the complete adequacy of the equipment presently in use, any change would very likely be uneconomical.

With reference to Mr. Levitsky's remark that his reference standards are operated continuously, our experience is that the self-heat errors of most watt-hour meters are reproducible, the asymtotic value of registration being reached within a few hours. Thus, it does not seem necessary to operate the meters continuously to obtain good precision. However, continuous operation is a convenience if tests are to be made every day.

This is an interesting account of one application of high-speed electronic counters to the testing of watt-hour meters. It should encourage the production of papers by others who are using electronic counters for the measurement or control of time, power, and energy in watt-hour meter testing.

G. N. Burkhart, Jr. (Duncan Electric Manufacturing Company, Lafayette, Ind.): Mr. Levitsky has brought out many points that should be of value to those planning the construction of test equipment. The general method discussed of indicating standard revolutions by photoelectric pickup and decade counting of revolutions and parts of revolutions is a modification that can be applied to any rotating standard. Most difficulties in rotating standards can be found either in the friction of the register or in improper operation of the resetting mechanism. The apparatus described by the author removes both of these items from the rotating standard, leaving a nearly friction-free device operating continuously without starting or stopping errors.

Similar circuitry to that described by Mr. Levitsky is in use at my company's plant. Recently a frequency doubling circuit has been added to the output of the photoelectric pickup to obtain twice as many bits of information from the standard, since this obtains a count on both the beginning and end of each light impulse.

It has been found possible to obtain a direct reading of meter accuracy on the counter by presetting the counter to 20,000 and running backward for a nominal 10,000 pulses. The result then is a number which for small percentage deviations from 100 is the actual percentage registration of the meter under test in one-hundredths.

Harold P. Knopp (Electrical Facilities, Inc., Oakland, Calif.): The author has made a worth-while contribution to the art of precision testing of portable watt-hour meter standards (rotating standards). Mr. Levitsky has taken, as he states, the best ideas which could be found among manufacturers and utilities and has added the new technique of employing two photoelectric systems, embodying in one a decimal counter for accurately measuring fractional turns without the customary register. Reference rotating standards have been employed previously without registers for the purpose of reducing friction error, but Mr. Levitsky's method makes it possible to test such standards by the definite-time-interval method. Therefore, at least in so far as equipment is concerned that employs the definite-time-interval method for checking the primary reference rotating standards, Mr. Levitsky's system should add appreciably to the accuracy and reliability of performance.

The author has recognized the increasing need for adequately accurate and dependable procedures and equipment for the testing of laboratory reference rotating standards. It is this problem that he attacks especially, and in his paper, when discussing his calibration procedure, he points out the fact that the reference rotating standards need be calibrated at only two points inasmuch as the standards are supplied from the secondaries of precision multirange transformers and therefore always operate with the same magnitude of voltage and current. From this it can be seen that the test procedure is simplified and that greater accuracy and dependability are obtained both in the calibration of the reference standards and in the utilization of these standards for field-standard testing. This is not all, however, and I feel that, especially in view of the author's particular concern with obtaining an accurate and dependable means of calibrating the laboratory reference rotating standards of his equipment, this matter is worthy of further discussion.

Actually, utilization, as in the Levitsky equipment, of precision multirange transformers in the arrangement just described, known as the Knopp uniload system,[1] produces benefits in the form of simplifications and improved accuracy and dependability that reach all the way down to the basic calibration at the d-c potentiometer level. This is true because the uniload system makes possible, when calibrating the laboratory reference rotating standards against the reference wattmeter, the use of a single-range indicating wattmeter which need be used only at two scale points. This is in place of several wattmeters to cover all ranges, the accuracy of each of which would have to be determined and maintained at many points on the scale. This simplification reduces to a minimum the need for expensive auxiliary standardizing equipment for use with the potentiometer, such as the precision shunts. Furthermore, laborious, time-consuming tests for determining and maintaining the accuracy of the various indicating wattmeters at many scale points are eliminated.

In the interest of dependable results Mr. Levitsky has devised techniques for intercomparing the standards and cross-checking the counters, and weekly tests on the reference standards of the equipment are made against the indicating wattmeter. The wattmeter connected as shown in Fig. 1 does not, however, check the accuracy of

Vol
Pt.
195

performance of the equipment as a whole, and it would be interesting to hear whether the author has made over-all checks and, if so, what results were obtained. I have reference to the need for calibrations that take into account not only the performance of the reference standards but also the performance of the precision transformers and the resistance of circuit components such as the switches in the potential circuit to the field standards under test. It is realized that the error of the precision transformers is negligible and that their performance is remarkably stable, but the potential switches could introduce errors caused by faulty or worn contacts. So as to check the performance of the equipment over-all, we recommend connecting the wattmeter to the 5-ampere and 120-volt ranges of the circuit to the rotating standard under test. Because of the very high interrange accuracy of the Knopp precision multirange transformers,[1] this calibration on the one current and one voltage takes care of all ranges. Should it be desired to verify, independently, the performance of the precision transformers, this can be done with necessary connections and without removing the transformers from the equipment by use of the Knopp one-to-one method.[1,2]

REFERENCES

1. See reference 1 of discussion by Mr. Pritchett.

2. COMMERCIAL STANDARDIZATION OF INSTRUMENT TRANSFORMERS, O. A. Knopp. *Electrical World*, New York, N. Y., vol. 67, Jan. 8, 1916, pp. 92–93.

F. J. Levitsky: I wish to thank, the discussers for their interest in this paper. The majority of the discussers particularly mentioned errors in the measurement of partial revolutions of a meter disk caused by the variation in speed of a disk during one revolution under constant load. Mr. Spencer points out, however, that the magnitude of this error will decrease as the number of full revolutions checked is increased, and with a sufficiently large number of revolutions the error would be negligible. Furthermore, this error is a ± quantity when tests are started at random as they are with this equipment. Because of this, and as stated in the paper, the possible ± 1 count error on the decade, counter, no single test result on the reference standards is ever used. The average of a large number of tests is used, and it is believed that these errors are of negligible magnitude in the average.

Mr. Austin described difficulties encountered in using a watt regulator caused by wave-form distortion. Tests made in the current circuit on the test board described in this paper with the watt regulator in the circuit showed a 5% second harmonic and a 1% third harmonic. There was no difference in harmonic content in the potential circuit with and without the watt regulator concerned. Calibrations made manually and with the automatic watt regulator when the test board was first placed in operation and at later dates showed no differences in performance values for the reference standards. As a result of these tests, it is assumed that the measured harmonics in these circuits caused by the watt regulator have negligible effect on reference standard performance.

Mr. Pritchett is correct in noting a missing connection in the schematic diagram shown in Fig. 1. He also mentions the fact that errors would be caused by short-circuiting the wattmeter current coil, thus changing the burden on the precision current transformers. This is true, but in practice the current coil of the wattmeter is left in the circuit during all tests, so that burden on the precision current transformer is constant and highest accuracy obtained.

---

ol. 74
. 1
55

A Fully Selective Telemetering System Employing Telegraph Facilities

C. W. SMITH
ASSOCIATE MEMBER AIEE

M. E. FORREST, JR.
NONMEMBER AIEE

IN APRIL of 1953 the Mississippi Valley Gas Company, whose headquarters are located in Jackson, Miss., approached commercial representatives of the Southern Bell Telephone and Telegraph Company concerning the provision of a telemetering system between their dispatcher's office in Jackson and 17 points in the central and northwestern section of the state to which natural gas was being furnished through their distribution lines. Up until this time individual meters located at these outlying points were read by employees stationed in the general area and the readings were telephoned to the dispatcher in Jackson. While this procedure was originally acceptably satisfactory, a steady growth in the number of residential customers and industries using gas was making the distribution problem progressively more acute. The more frequent readings of the outlying meters required not only increased the cost of the meter reading but necessitated the adoption of a more speedy and efficient system. It was proposed that such a system be obtained by the use of telemetering equipment operating over telegraph facilities.

The use of economical telegraph channels for the remote control of equipment or reading of meters has proved advantageous in many cases. Where continuous readings are not required, equipment to permit the same line to be used to read a large number of meters or control a large number of devices, one at a time, will affect large economies in telegraph channel mileage.

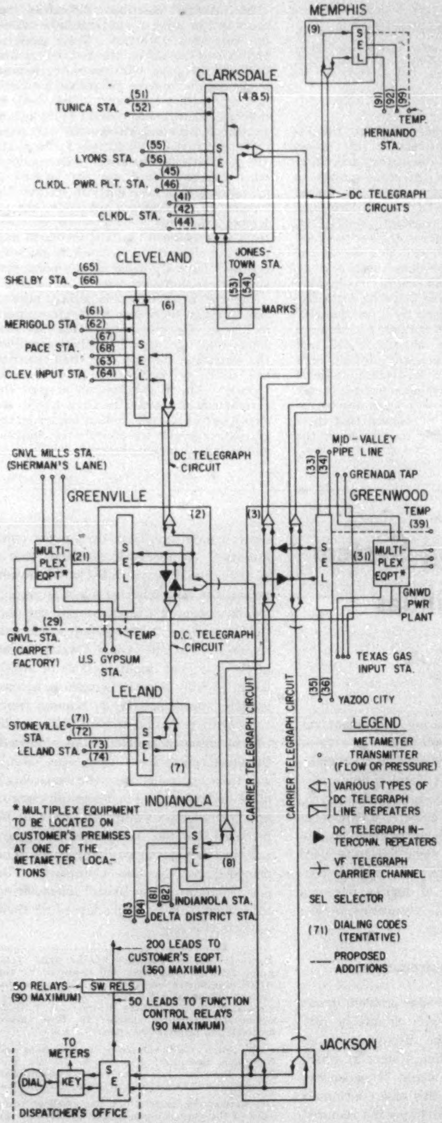
Basic Design Requirements

To meet this particular problem it was proposed that telegraph circuits be provided to connect the dispatcher to all points on the pipe line system at which readings or control would be necessary. It was felt that in this case continuous readings from all points were not required and would entail far too great an expense for circuit mileage. Automatic sequential readings did not appear to meet fully the requirements because more frequent readings were needed from the major outlet points than from the smaller communities. It appeared, therefore, that a system should be provided which would permit the dispatcher to select any point on the system at will and obtain the particular reading at that point which was required. No valve control was required at the start but might need to be added later. A fully selective system to accommodate approximately 50 reading functions with provisions for the later addition of valve control functions appeared, therefore, necessary. The entire circuit was to be controlled from the dispatcher's office in Jackson. The telegraph circuits and the selective equipment were to be provided by the Telephone Company, and the telemetering equipment was to be provided by the Gas Company. The Gas Company chose Bristol telemetering equipment as being well adapted for their needs in this case.

Paper 55-167, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 26, 1954; made available for printing December 9, 1954.

C. W. SMITH is with the American Telephone and Telegraph Company, New York, N. Y., and M. E. FORREST, Jr., is with the Southern Bell Telephone and Telegraph Company, Atlanta, Ga.

The authors are grateful for the excellent team work of the various organisations involved in the provision of this system. The co-operation of L. B. Lumpkin of the Bristol Company, Birmingham, Ala., and W. E. Rufleth, their Systems Development Engineer at Waterbury, Conn., with the personnel of the Southern Bell Telephone and Telegraph Company resulted in the successful completion of the installation in the minimum time.


MEMPHIS
(9)
SEL
(91)
(92)
(66)
TEMP.
HERNANDO STA.
CLARKSDALE
(4&5)
TUNICA STA.
(51)
(52)
LYONS STA.
(55)
(56)
CLKDL. PWR. PLT. STA.
(45)
(46)
CLKDL. STA.
(41)
(42)
(43)
(44)
DC TELEGRAPH CIRCUITS
JONES-TOWN STA.
(53)
(54)
MARKS
CLEVELAND
SHELBY STA.
(65)
(66)
(6)
MERIGOLD STA
(61)
(62)
PACE STA.
(67)
(68)
CLEV. INPUT STA.
(63)
(64)
DC TELEGRAPH CIRCUIT
GNVL MILLS STA. (SHERMAN'S LANE)
GREENVILLE
(2)
MULTI-PLEX EQPT*
(21)
(29)
GNVL. STA. (CARPET FACTORY)
TEMP
U.S. GYPSUM STA.
D.C. TELEGRAPH CIRCUIT
LELAND
STONEVILLE STA.
(71)
(72)
LELAND STA.
(73)
(74)
(7)
INDIANOLA
(8)
(81)
(82)
(83)
(84)
INDIANOLA STA.
DELTA DISTRICT STA.
* MULTIPLEX EQUIPMENT TO BE LOCATED ON CUSTOMER'S PREMISES AT ONE OF THE METAMETER LOCA-TIONS
MID-VALLEY PIPE LINE
(33)
(34)
GRENADA TAP
GREENWOOD
TEMP
(39)
(3)
(31)
MULTI-PLEX EQPT*
GNWD PWR PLANT
TEXAS GAS INPUT STA.
(35)
(36)
YAZOO CITY
CARRIER TELEGRAPH CIRCUIT
CARRIER TELEGRAPH CIRCUIT
LEGEND
METAMETER TRANSMITTER (FLOW OR PRESSURE)
VARIOUS TYPES OF DC TELEGRAPH LINE REPEATERS
DC TELEGRAPH IN-TERCONN. REPEATERS
VF TELEGRAPH CARRIER CHANNEL
SEL SELECTOR
(71) DIALING CODES (TENTATIVE)
PROPOSED ADDITIONS
200 LEADS TO CUSTOMER'S EQPT. (360 MAXIMUM)
50 RELAYS (90 MAXIMUM)
SW. RELS.
50 LEADS TO FUNCTION CONTROL RELAYS (90 MAXIMUM)
TO METERS
DIAL
KEY
SEL
DISPATCHER'S OFFICE
JACKSON

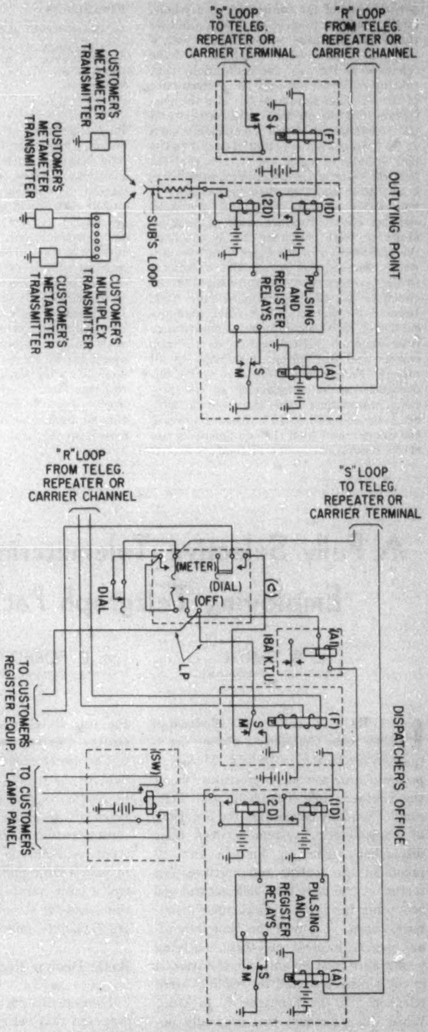

"S" LOOP TO TELEG. REPEATER OR CARRIER TERMINAL
"R" LOOP FROM TELEG REPEATER OR CARRIER CHANNEL
CUSTOMER'S METAMETER TRANSMITTER
CUSTOMER'S METAMETER TRANSMITTER
CUSTOMER'S METAMETER TRANSMITTER
CUSTOMER'S MULTIPLEX TRANSMITTER
SUB'S LOOP
S
M
(F)
(2D)
(1D)
PULSING AND REGISTER RELAYS
M
S
(A)
OUTLYING POINT
"R" LOOP FROM TELEG. REPEATER OR CARRIER CHANNEL
"S" LOOP TO TELEG. REPEATER OR CARRIER TERMINAL
(METER)
(DIAL)
(OFF)
(C)
DIAL
LP
(AI)
18A KTU
TO CUSTOMER'S REGISTER EQUIP.
S
M
(F)
(SW)
TO CUSTOMER'S LAMP PANEL
(2D)
(1D)
PULSING AND REGISTER RELAYS
M
S
(A)
DISPATCHER'S OFFICE

Fig. 1 (left). **Over-all telegraph circuit layout for selective telemetering, Mississippi Valley Gas Company**

Fig. 2 (above). **Application schematic of telemetering selector unit**

Physical and Geographical Considerations

The first appraisal of the proposed circuit layout indicated the installation of selector equipment at each of the outlying points, with the possibility that the equipment would have to be mounted on a pole, or housed in a small corrugated sheet iron shack at each location. Since only three functions (pressure, flow, and temperature) would be read at any one location, the proposed arrangement appeared to involve an extremely extravagant use of selectors. Further, the necessity of installing such equipment, together with telegraph repeater equipment, on poles or in shacks appeared undesirable.

Geographically, the majority of the 17 points from which readings were desired are located in the delta section of the state. According to Weather Bureau records the temperature variations in this area ranged from 12 degrees below zero to 110 degrees above, Fahrenheit, over a 40-year period. Relative humidity readings taken at 6:30 a.m. daily during the months from May through September indicated a range from 77 to 90 per cent. It was also known that dust from the silt deposits in the area penetrated all but waterproof housings and that innumerable varieties of insects infested any exterior housing offering access. Cotton plantations extend for miles on end throughout the section and the spraying, or dusting, of these fields by airplanes is common practice. Some of the chemicals used in these methods of pest control leave a gummy deposit on exposed metal and initiate a corrosive action. This combination of conditions offered almost every conceivable hazard to the satisfactory operation of equipment located in other than heated and regularly occupied buildings.

An investigation was made of the possibility of using a circuit layout which would permit the selection of a number of functions at several points in the same general locality from a single selector. A layout was developed on this basis which would concentrate the desired functions in seven selectors which could be installed in established telephone central offices. The circuit layout is shown in Fig. 1. The arrangement required the extension of a number of the function legs (the circuit between the selector and the Bristol transmitter equipment) for moderate distances beyond the offices in which the selectors would be installed, but the overall redesign effected a net saving to the customer for monthly rentals of equipment and line facilities.

Circuit Development Work

The next part of the problem was to choose a selector adaptable to standard telegraph line facilities and repeater equipment, and compatible with the Bristol impulse duration system of metering. There appeared to be no available selector which was in all respects satisfactory from an operating standpoint and which was sufficiently rugged to give continuous trouble-free service at isolated locations for long periods of time.

As only a short interval was available for development work, manufacture, and installation, investigation was directed toward some sort of basic selecting arrangement which would provide the desired features, the component parts of which could be obtained immediately from supplier's stock. Consideration was given to the use of stepping switches and various types of mechanical selectors, but these were rejected as presenting a maintenance problem when installed at outlying points. The most promising solution appeared to be an all-relay arrangement capable of responding to dial pulses and, in accordance with the dialed code, able to effect connection to the line of one of any desired number of telemetering devices. Similar arrangements are widely used in the telephone plant in connection with dial operation. One such circuit for responding to dial pulses was described by J. W. Dehn.[1]

The selector circuit shown schematically in Fig. 2 consists of several functional parts:

1. A polar receiving relay to receive the dial pulses from the line and repeat them into the selector.
2. Two slow release relays, one to operate when the line is seized for dialing, and remain operated until the selector is dismissed, the other to operate on the first dial pulse and remain operated during the train of pulses composing a digit, but release between digits.
3. A pulse-dividing circuit to feed the dial pulses to the counting relay chain.
4. A counting relay chain to count the pulses in each digit.
5. A relay for each first digit of the code to which the selector is intended to respond.
6. Interdigital relays to switch to second-digit relays.
7. A second-digit relay for each second digit of the code to which the selector is intended to respond.
8. A pulse-absorbing arrangement to absorb any first digit 1 which might be received to prevent false selection as a result of line hits.
9. A connection through the contacts of operated first-digit and operated second-digit relays to perform the required switching function.

Since the first digit 1 was eliminated as a valid code, the selector could be arranged to respond to a maximum of 90 2-digit codes. By adding third-digit relays and the necessary interdigital relays, the selector could readily be expanded to respond to 900 3-digit codes. A selector could be equipped with all the digital relays and, therefore, be capable of responding to all the codes, or it might be equipped with only a portion of the digital relays and so be capable of responding to only a few of the possible codes.

The output circuit also was equipped with a polar relay *F*, Fig. 2, arranged to connect to the loop circuit of any standard telegraph repeater or telegraph carrier channel. Since the coded function leg terminals in each of the outlying selectors would generally be extended to the customer's equipment over single wires of different resistances, a current-adjusting resistance was provided for cross connection into each of the function legs. The *225A* output relays *F* were arranged for a nominal 0.020-, or 0.060-ampere loop current.

To supervise the connection of meters or registers in the dispatcher's office for the reading of the intelligence transmitted from the outlying instruments, a selector was installed in the control room on the customer's premises in Jackson as shown in the right-hand side of Fig. 2. This selector is basically the same as those in the outlying offices. The input circuit *A* relay was connected in series with the

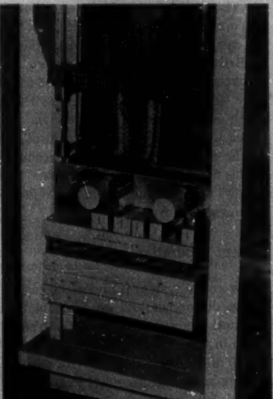


Fig. 3. Selector unit as installed in central office at Greenwood, Miss.

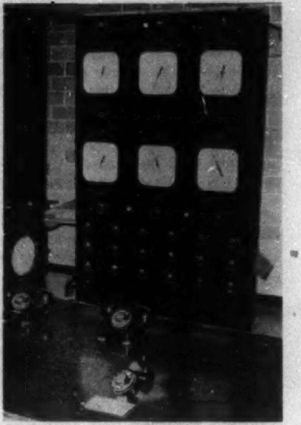


Fig. 4. Bristol meter and lamp panel in dispatcher's office at Jackson, Miss. Control position appears in foreground

line so that this selector responded to all codes dialed by the dispatcher. One purpose of this selector was to light an identifying lamp for each function dialed for checking purposes. Provision was also made at the request of the Bristol Company to arrange the selector for the operation of two additional devices for each function dialed. Therefore, 50 switching relays *SW* were installed and arranged to permit the cross connection of any relay to any one of the 90 coded line terminals. The armature and contact circuits of these relays were wired to a demarcation terminal block to permit the customer to handle the connection of his Bristol apparatus without making any soldered connections on Telephone Company equipment. The metering leads, which are extensions of the receiving leg of the line circuit through a control key on the dispatcher's desk, also terminate on this demarcation terminal block. Fig. 3 shows a selector unit as installed in the Greenwood central office.

System Operation

The entire system is under the control of the dispatcher and on his table are mounted the control key *C*, Fig. 2, referred to in the foregoing, key position pilot lamps, and a dial. The control position appears in the foreground of Fig. 4. The control key has three positions, designated OFF, DIAL, and METER. Normally, the key remains in the OFF position which gives a spacing condition on the outward leg of the circuit. To seize the circuit and prepare the selectors for receiving the dial pulses, the key is operated to the DIAL position which operates the outward circuit leg to the marking condition. The two code digits for the desired reading are then dialed and all selectors follow the dial pulses but a function leg is connected to the circuit at only one outlying selector, the one conditioned to respond to the particular code dialed. The selector in the control office also responds and operates the switching relay *SW* associated with the dialed code. The contacts of this relay control the lamp signals and the connection of the proper meter in the customer's display panel. Upon the operation of the control key to the METER position, the outward, or dialing, leg of the circuit is held closed to maintain the selector positions and the inward leg of the circuit is connected to the METER leads so as to direct the incoming telemetering pulses to the proper meter for registering, or recording. The pilot lamps on the table serve to notify the dispatcher when the circuit is being held in an operative condition. When the reading has been completed, the control key is restored to normal which opens the telegraph sending leg and releases all selectors so that they are ready for another call.

Application to Telegraph Line Sections

The telegraph lines used with the system are all arranged with separate send and receive legs as for full duplex operation, although transmission is in only one direction at a time, the other direction being held steady marking. At the dispatcher's office the dial pulses are sent out on the sending leg and received on a selector on the receiving leg of each station on the line. At the selected outlying station the telemeter sending device is connected to the sending leg and, at the dispatcher's office, the telemeter reading device is connected to the receiving leg.

Fig. 5. Front view of customer's installation of Bristol metameter transmitters (right), at Greenwood input station

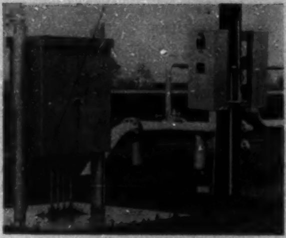


Fig. 6. Rear view of Greenwood input station installation

With the full duplex arrangement, the slow telemetering pulses are prevented from appearing on the receiving leg of the repeater into which they are sent so that they can not interfere with the selector connected to that receiving leg. These pulses are likewise prevented from appearing on any other leg on the system containing a selector. Even if telemetering pulses should be received at a selector other than the one at the point where they originate, they would be interpreted as a series of line hits and would be absorbed by the selector and thus do no harm.

Another feature of interest is that the *F* relay in the outlying selectors is biased to marking instead of to spacing. This marking bias serves to hold the sending leg of the circuit to marking when no telemtering pulses are being transmitted. Thus, the send relay in the d-c repeater is in its normal marking position, as it should be, while dial pulses are coming in on the receiving leg. The telemetering pulses operate the *F* relay to spacing.

Arrangement of Telemetering Instruments

A multiplex transmitter for the telemetering readings was provided at Greenwood and at Greenville, each arranged to transmit seven readings in sequence when the proper code for the multiplex unit was dialed. This change was considered necessary since the volume of gas to be drawn from the distribution system at these points is so great that frequent readings are essential. With the multiplex arrangement, the dispatcher by dialing one 2-digit code can secure in sequence the seven readings from the seven metameter transmitters connected to the multiplex transmitter with a resultant saving in time. The arrangement also necessitated the provision of a multiplex receiver at the dispatcher's office.

The multiplex equipment was provided with an arrangement which, upon the closure of the multiplex leg by dialing its code, synchronized the multiplex transmitter with the receiver and immediately began the sequence. Without this arrangement, there could be a lengthy delay between the time the transmitter was connected and the sending and receiving ends were brought in synchronization. Figs. 5 and 6 show the customer's installation at one of the points out of Greenwood, from which point three functions are read through the multiplex equipment. This arrangement is typical of the installations at the other metering points.

Code Call Assignments

The assignment of codes to the various points was made, as indicated in Fig. 1, to simplify the trouble-reporting procedures. By assigning one, or two, groups of ten codes to each selector, the customer finds it relatively simple to check all codes operative through any one selector and, if trouble is experienced, report that no code in a particular group can be registered. Telephone Company plant operating personnel also find this grouping helpful in isolating, or sectionalizing, reported troubles.

Preservice Tests

On the initial line-up tests made just prior to placing the system in service, telegraph transmission measurements were made between Jackson and each selector office on the outward circuit leg. Tests also were made from the outlying end of each function leg while it was selected to the Jackson office over the inward circuit leg. All transmission tests indicated the over-all circuit to be suitable for 60-speed teletypewriter signals which are 22 milliseconds in length. Related to the telemetering pulses, which range from 1.0 to 4.0 seconds in length, any metering error resulting from telegraph distortion would be so small as to be negligible.

Maintenance Experience

Maintenance attention on the telegraph circuit and selector equipment has been normal since it was placed in service and the Gas Company has reported the service to be entirely satisfactory.

Conclusion

The solution to the problem outlined was worked out with a view to meeting an urgent request for service where existing instrumentalities were, for one reason or another, not applicable and where ease of operation and freedom from service interruptions were of the highest importance. The success of this system for the Mississippi Valley Gas Company has been such that it may well be employed in other cases where it appears to meet the requirements.

Reference

1. The Original Dial Pulse Register Circuit for the Crossbar System No. 5, J. W. Dehn. *Bell Laboratories Record*, New York, N. Y., Jan. 1950.

◆

Discussion

W. E. Rufleth (The Bristol Company, Waterbury, Conn.): The authors should be complimented for a most excellent paper. I believe the Mississippi Valley Gas Company installation is the first gas-dispatching system using this particular type of selector furnished by a communication company, and it opens up a vast field for this type of application.

To elaborate further on the dispatcher's identification lamps at the Jackson Office, it might be said that, when a particular telemeter transmitter is being contacted, a lamp is lighted to show the metering station involved. Another lamp is lighted to indicate whether the reading is pressure, flow, temperature, or multiplex system contacted, and still another lamp on the telemetering receiver indicates on what particular receiver the signal is coming in.

Possibly Mr. Smith and Mr. Forrest are acquainted with the selective equipment used by the Tennessee Gas Transmission Company near Clendenin, W. Va., and furnished by Chesapeake and Potomac Telephone Company and the American Telephone and Telegraph Company. Will they please mention whether it would be possible to actuate graphic-panel lights from read-back signals in connection with the selective system described in their paper? In other words, at present, if the dispatcher at Jackson, Miss., wishes to read the Hernando station static pressure, he deals 91 while holding the control key in the DIAL position. The selector in the dispatch office also responds to the dialed 91 and lights a signal light indicating that the number for obtaining the static pressure at Hernando station has been dialed. Would it be possible, without too much complication, to have a read-back signal come back from Hernando's to indicate that the Hernando's telemetering pressure transmitter has been contacted?

Samuel A. Scharff (Systems Engineer, New York, N. Y.): This interesting description of a telemetering system arouses the appetite for more, and it is hoped that the authors will give quantitative data on the reliability of the system in future papers. The scope of the present paper could not include complete data on reliability. My impression is that recorded data on system reliability are rare.

I have reviewed 50 items from my files. All deal with remote control, supervisory control, and telemetering, not all from the field of public utility systems. All were published within the last 10 years. They do not, however, represent a comprehensive file on the subject.

Of the 50, only two give specific, quantitative records of performance.[1,3] W. F. Clare's paper[1] is unique in its comprehensive account, based on 2 years' observations, of incorrect control operations, incorrect indications, failures to operate, and failures to indicate in several types of equipment used by a large electric power company.

One paper reported "...no repairs or replacement...only occasional minor adjustment..." from 1 year's use of a small system (apparently 1-channel, possibly multiplexed for four quantities). Two report "...reliable and consistent..." operation or similar words reflecting qualitative satisfaction; but this provides no basis for estimating the usefulness of the equipment or approach in other applications.

H. H. Joyner's paper[3] is a most important contribution to the subject since it reports reliability of the American Telephone and Telegraph Company leased channels, but it naturally includes no discussion of the telemetering equipment. An AIEE Committee Report[4] states that, of 75 repliers, 79 per cent consider automatic or supervisory controlled stations as reliable as manually operated stations. This is a useful report, of course, nevertheless it has a tantalizing side.

The Bibliography on Telemetering[5] which lists about 750 items with dates later than 1920 contains not one with the word "reliability" in its title. This, of course, is not a highly significant fact; neither does Mr. Clare's paper.[1] Yet perhaps it is not wholly irrelevant.

The other 42 papers reviewed contain no data on reliability. Some discuss possible sources of error and present results of tests of the effects of such errors; but no operating experience is recorded. Since the sample was collected almost at random, significant papers may well have been omitted. However, this brief review strengthens the position that data on reliability beyond "...serves admirably..." would be welcome.

Both manufacturers and using organizations are in positions to obtain information of this sort. Unsatisfactory performance may be thought painful to report; but users who are proud of their operating records and manufacturers with real engineering confidence in their products need not hesitate. They can reasonably expect that the reporting of solid experimental data would bring credit upon themselves, and it would certainly be of value to the profession.

REFERENCES

1. SUPERVISORY CONTROL'S CONTRIBUTION TO IMPROVED SYSTEM OPERATION, **W. F. Clare.** *Electric Light and Power*, Chicago, Ill., Oct. 1954, pp. 113–17.

2. A RADIÓ REMOTE-CONTROL SYSTEM FOR FM BROADCAST STATIONS, **T. R. Humphrey.** *Electrical Engineering*, vol 73, April 1954, pp. 327–30.

3. LEASED CHANNELS FOR METERING AND CONTROL, **H. H. Joyner.** *Electric Light and Power*, Chicago, Ill., Dec. 1953, pp. 100–06.

4. SUPERVISORY CONTROL AND ASSOCIATED TELEMETERING EQUIPMENT—A SURVEY OF CURRENT PRACTICE, **AIEE Committee Report.** *AIEE Transactions*, vol. 74, pt. III, April 1955, pp. 36–38.

5. BIBLIOGRAPHY ON TELEMETERING, **C. A. Mabey.** *AIEE Special Publication S-68*, Dec. 1954.

C. W. Smith and **M. E. Forrest, Jr.**: We wish to thank Mr. Rufleth and Mr. Scharff for their discussions. Regarding the provision of a read-back signal, or coded answer-back to actuate the signal lights at the dispatcher's office, and thus afford assurance that the proper function has actually been selected, the telemetering system of the Mississippi Valley Gas Company, as described in the paper, does not have this feature. It could be provided, however, by the addition of pulse-generating and coding equipment at each of the outlying points where selector equipment is provided. In the particular application discussed, the provision of this feature was considered, but the cost of the additional equipment required was thought not to be justified.

We agree with Mr. Scharff's comment to the effect that quantitive data regarding the performance of telemetering systems would be highly desirable. The absence of such data regarding the performance of the telemetering system described in the paper was not due to any lack of interest in this vital point, but because of insufficient man-hours for keeping an accurate performance record.

Shielding of Communication Cables

F. H. GOODING
MEMBER AIEE

H. B. SLADE
ASSOCIATE MEMBER AIEE

IN THE course of designing and manufacturing electric cables, the need for shielding frequently arises, sometimes because of inductive interference and sometimes because of high-voltage considerations. This discussion is confined to relatively low-voltage cables where inductive interference is the problem.

In communication receiving circuits where the amplification of extremely small signals is possible, the ratio of the received signal to the received noise is more significant than the absolute value of either the signal or the noise. It follows therefore that, provided sufficient amplification is made available, the length of a circuit will be limited by the total noise, a large part of which may derive from induction in the cable circuit or open-wire line connecting the transmitter to the receiver.

In open-wire lines noise may be considerably reduced by transposing the conductors, which assists in equalizing the induced voltages in the two sides of the circuit. These are in opposition in their direction of flow around the circuit and tend to balance out. In a cable circuit consisting of a twisted pair the conductors are continually transposed, and noise due to electrical disturbances in the vicinity may be further reduced by surrounding the cable circuit with a conducting or semiconducting shield. In a coaxial cable, the outer conductor is also the shield. It is impractical to transpose the inner and outer conductors. Surrounding the outer conductor by elaborate shielding often becomes necessary.

A shield may operate in three more or less distinct ways to reduce the field strength within the shield enclosure:

1. A continuous metallic shield will completely screen the electric field from circuits within the shield, for all practical purposes. "Electric field" is used in preference to "electrostatic" or "capacitive," and refers only to the high-impedance field associated with capacitive coupling.

2. A shield made of magnetic material will divert the magnetic field away from the conductors inside the shield.

3. Any metallic shield may reduce the magnetic field within it because of eddy currents set up within the material of the shield when exposed to an alternating magnetic field. This is because the currents set up fields which tend to neutralize the disturbing field.

At frequencies up to 2 or 3 megacycles (mc) it is convenient to study the effects of electrical disturbances by dividing them into electric field and magnetic field effects. The electric field disturbances are a function of the capacitance coupling or mutual capacitances between the disturbing and the disturbed circuit, while the magnetic field disturbances are proportional to the magnetic coupling or mutual inductance between the circuits. At higher frequencies energy is emitted from the disturbing source in the form of radiation which increases as the square of the frequency, and becomes a much greater factor in causing disturbances at considerable distances than the electric or the magnetic fields. This electromagnetic penetrates the cable shielding through interstices in braids or metal tape wrappings, as well as passing, with a certain amount of attenuation, through the material of the shield itself.

Shielding of Twisted Pair Cable

The effectiveness of the shielding applied over a twisted pair cable may be measured more directly than the shielding which forms the outer conductor of a coaxial cable, for in the latter case the shielding forms part of the communication circuit. To evaluate shielding over a 2-conductor cable circuit, a 10-foot length of flat-pair television cable was stretched on a board, and a disturbing circuit consisting of two parallel straight wires was laid parallel to the cable, spaced about 1 inch from it, all wires being in the same horizontal plane as in Fig. 1 (A and B). The disturbing circuit *AB* was energized from an oscillator and the induced voltage in the disturbed circuit *CD* was measured with a vacuum-tube voltmeter. A flat

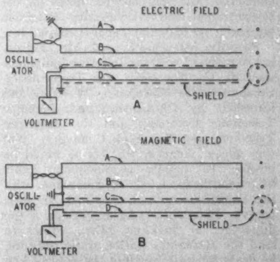
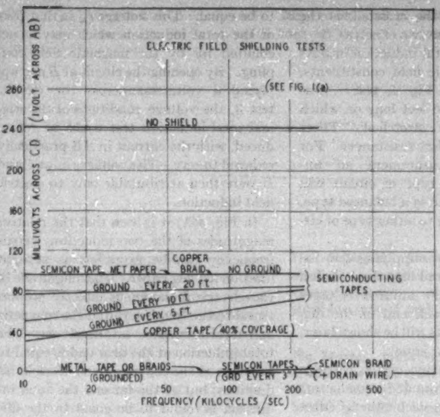


Fig. 1. Shielding test circuit for measuring induction. A—Due to electric field. B—Due to magnetic field

---

Paper 55-198, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted July 29, 1953; made available for printing December 2, 1954.

F. H. GOODING and H. B. SLADE are with The Okonite Company, Passaic, N. J.



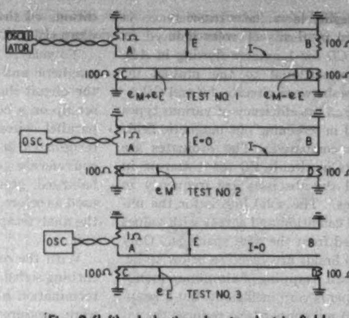


Fig. 2 (left). Induction due to electric field

Fig. 4 (above). Test circuits with terminated lines to analyze inductive components

cable was used for the disturbed circuit because there was then sufficient induction in a short length of cable to obtain readable indications on a low-range vacuum-tube voltmeter. This nontransposed arrangement is considered to duplicate the lack of uniform twisting in a long length of, say, 1,000 feet of twisted pair cable. The induction in the cable conductor nearest to the disturbing source was greater than in the more distant one. The vacuum-tube voltmeter measured the difference in the induced voltages. Tests were first made with the cable pair *CD* unshielded, and subsequent tests with various types of shielding over the pair. Effectiveness of the shield or shielding factor may be obtained from the ratio of induced voltage in the shielded circuit to induced voltage in the unshielded circuit.

Electric Field Shielding

This is a function of the capacitance coupling between the *AB* and *CD* circuits. To eliminate magnetic coupling between circuits, the wires were open at the far end so as to reduce the current flowing in them practically to zero, and grounds were applied to the outer wire of each circuit, Fig. 1(A). Interference was measured in terms of millivolts induced in *CD* for 1 volt across *AB*. When metallic shielding was applied over *CD*, it completely eliminated any coupling between the two circuits with the shield grounded, and there was therefore no reading on the voltmeter. The voltages induced in circuit *CD* for various shields and ground conditions are plotted in Fig. 2.

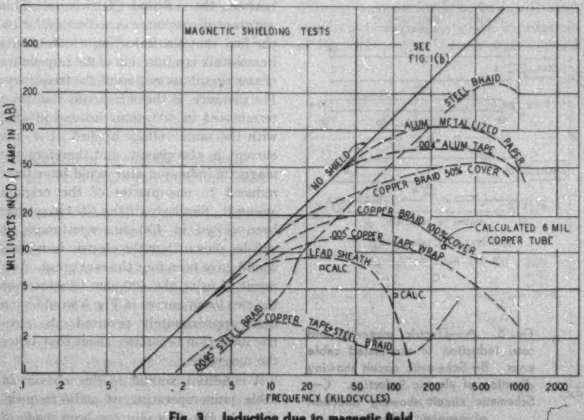


Fig. 3. Induction due to magnetic field

An interesting possibility is introduced where magnetic shielding is not required because of small currents or low frequencies in the disturbing circuit, by using a semiconducting shield in place of a metallic shield. This semiconducting shield may be highly efficient as an electric field screen and at the same time much less expensive. For this reason tests were made with semiconducting tapes and braids, the braids consisting of semiconducting yarns which were lighter, and more flexible than metal braids. Because of their higher longitudinal resistance, it is necessary to insert a single bare "drain wire" into the cable assembly which contacts the covering, and grounds it along its length. Fig. 2 illustrates the effectiveness of grounding at various intervals. It will be seen that while grounds every 10 or 20 feet will improve the electric-field shielding, it is desirable to ground the shield with a drain wire which contacts the semiconducting shield at least every few inches.

An ungrounded shield will reduce the induction in *CD* to less than half the value for no shield, and a metal shield occupying 40 per cent of the surface shields out three-quarters of the induction.

Magnetic Field Shielding

To measure the value of various types of shield in screening out electromagnetic fields, the circuit was set up as in Fig. 1 (B) with the two circuits short-circuited at the far end. By grounding wires *B* and *C* electrostatic coupling was reduced to a negligible value. Interference is expressed in terms of volts induced in circuit *CD* for 1 ampere flowing in *AB*, and is proportional to the mutual inductance between circuits *AB* and *CD*.

In Fig. 3 the efficiency of various types of shield in screening out magnetic fields may be compared. The ordinates are millivolts in circuit *CD* for 1 ampere in *AB* and the abscissas are frequency in kilocycles. The solid line is for the unshielded condition and agrees with values calculated from the wire spacings. Only the steel braids are effective below 10 kc. At 100 kc a wrapping of aluminum metallized paper is very inefficient, but a wrapping of 5-mil copper tape reduces the induced voltage to one tenth of that for unshielded cable. A lead sheath is still more effective, and the figures for lead agree well with values calculated from an equation given by Terman.[1] Copper braids are fairly efficient above 100 kc, but a combination of copper tape (or braid) and steel braid is effective even at audio-frequencies. Steel conduit, not shown on the graph, is the most effective of all.

Induction Between Cable Pairs Terminated with Impedances

Having established how inductive interference or crosstalk may be treated as two more or less distinct phenomena, magnetic and electric, although they are fundamentally identical, it is necessary to examine the practical case of a circuit which is not open or closed, but which is terminated in an impedance, generally equal to its own characteristic impedance. In this case the current and voltage conditions coexist and the crosstalk is the phasorial sum of these two effects.

To analyze the total induction into its magnetic and electric field constituents, the circuit shown in Fig. 4, test 1, was set up on a board 10 feet long on which parallel wires were stretched. These terminated in 100-ohm resistances. For convenience of measurement an unbalanced, grounded type of circuit was used as before, instead of a balanced type, the analysis applying to either type of circuit.

With the oscillator supplying the disturbing signal at *A* and into the 100-ohm termination at *B*, the induced voltages were measured at each end of the disturbed circuit *CD*. As will be shown later, these voltages are not equal.

Most of the induction occurs between the adjacent ungrounded conductors which are nearest to each other. After making the total induction tests, the current *I* in the circuit was obtained (calculated from the voltage across *B*), and the voltage measurement *E* across *B* gave the voltage of the "high" side of the *AB* circuit above ground. The magnetic induction e_M in *CD* is a direct result of the current in *AB*, while the electric induction e_E is a function of the voltage of the high side of *AB* above ground. The total induction at *C* is e_M+e_E.

To determine the magnetic induction e_M independently, the far end of circuit *AB* was then short-circuited in test 2 and the current in *AB* adjusted to have the same value *I* as in the previous test. A voltmeter applied across a 1-ohm resistor indicated the value of the current. The voltages across *C* and *D* were then found to be equal. This voltage e_M is that part of the total induction which may be accounted for by the magnetic field coupling. By opening the circuit at *B* and applying a voltmeter across the circuit, test 3, the voltage conditions of the original total induction test could be reproduced, with the current in *AB* practically reduced to zero. The voltages e_E at *C* and *D* were then attributable only to electric field induction.

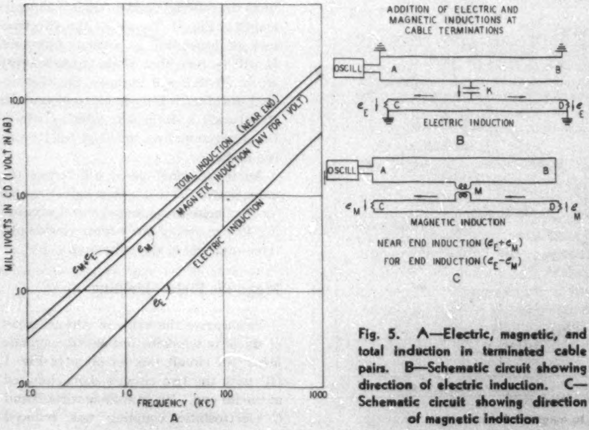


Fig. 5. A—Electric, magnetic, and total induction in terminated cable pairs. B—Schematic circuit showing direction of electric induction. C—Schematic circuit showing direction of magnetic induction

In Fig. 5(A) it is seen that the relative magnitudes of the two inductions versus frequency can be expressed as straight lines on log/log paper. The magnitude of the electric induction is directly proportional to the frequency, and the magnetic and total inductions are nearly so. The total induction at the near end is equal to the sum of the magnetic and electric inductions, but at the far end the total induction is found to be equal to the difference of the two inductions. The reason is that if the total effective capacitive coupling between the *AB* and the *CD* circuits is represented by a capacitance *K* as shown in Fig. 5(B), the resultant capacity currents will flow in the directions shown by the arrows. The currents resulting from the mutual inductance *M*, Fig. 5(C) which couples the circuits together will be in opposite directions at the two ends of the circuit. Therefore, the total voltage at the near end is (e_E+e_M) and at the far end (e_E-e_M).

Over the whole frequency range the magnetic is four to six times the electric induction. This statement may appear to be at variance with the common experience that at audio-frequencies crosstalk due to induction may be attributed almost entirely to unbalance capacitances between the *AB* and *CD* circuits. This apparent discrepancy is accounted for by the fact that the induction which results in crosstalk is a function of the impedance of the circuit, as well as of the frequency. For instance, if the *AB* circuit had been terminated in 400 ohms instead of 100, with the same voltage applied at *A*, the current in the circuit, and therefore the magnetic induction also would have been reduced to one-quarter of the original amount. Similarly, if the *CD* circuit has been closed in 400-ohm resistances, it will be shown that the electric induction would have been four times as great. As a result of using the 400-ohm terminations the two lower curves in Fig. 5 would have been approximately reversed, the electric induction becoming about four times the magnetic.

A condition similar to this obtains in cable pairs operating at audio-frequencies. In Fig. 6, a plot has been made of

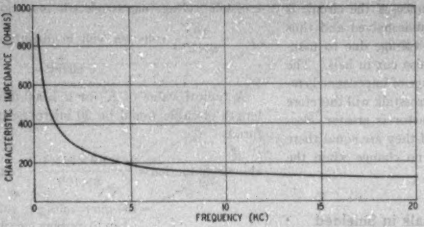


Fig. 6 (above). Characteristic impedance versus frequency of spiral-four type of cable

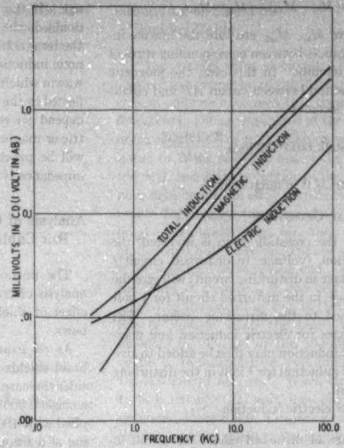


Fig. 7 (right). Inductive coupling between cable pairs as in Fig. 5 but terminated with long lengths of spiral-four type of cable

the characteristic impedance of the "spiral-four" type of cable, similar to that used by the United States Army Signal Corps, over the audio range. At radio frequencies the impedance of this cable is of the order of 130 ohms, but at 1 kc the impedance of a long line is over 400 ohms. This is explained by the relatively greater importance of the resistance of the cable at the lower frequencies. At high frequencies, the resistance is negligible compared to the inductive reactance, and the impedance is equal to the surge impedance L/C. But in the audio range R cannot be neglected, and the impedance is given by

$$Z=\sqrt{\frac{R+j\omega L}{j\omega c}}=\sqrt{\frac{L}{C}-j\frac{R}{\omega c}}$$

The effect of this increasing impedance may be estimated by inserting the 10-foot unbalanced circuit shown in Fig. 4 in the center of a long length of otherwise balanced spiral four cable. In other words, the flat wires of test 1, instead of being terminated in 100 ohms, would be terminated with long lengths of spiral-four cable at each end. Fig. 7 illustrates the reversal of position of the electric and magnetic crosstalk below 1,500 cycles, which occurs because of this increasing impedance at low frequencies.

Chapman[2] has shown how unbalance capacitances, inductances, and resistances produce noise currents in the Signal Corps spiral-four cable.

Calculation of Crosstalk Between Pairs in a Multipair or Quad Cable

A theoretical analysis of the unbalance conditions which result in inductive interference or crosstalk has been comprehensively given by Hayes, Seymour, and Bray.[3] It is shown how the electric and magnetic effects may be computed from measurements of the capacitances and inductances between wires.

Fig. 8(A) shows the capacitances between wires of balanced (to ground) pairs AB and CD neglecting the ground capacitances since we are considering side-to-side unbalance in which the ground capacitances are of less importance. The network may also be represented by the four impedances of Fig. 8(B). If the inductive effects between AC and AD are equal there will be no resultant voltage in circuit CD due to A and similarly for B. Or if the inductive effects of A and B on circuit CD are equal and opposite there will be no resultant voltage in CD due to AB. This is expressed by

$$K=(Z-W)-(Y-X)$$

It is shown in reference 3 that the crosstalk is proportional to the unbalance capacitance K. A further analysis shows that the voltage coupling between AB and CD may be illustrated by the capacitance couplings of Fig. 8(C) or the impedances of Fig. 8(D). If a voltage V is applied across AB, the voltage across CD will be in the proportion that impedance $Z/2$ bears to the total impedance across AB; see Fig. 8(D). This gives

$$\text{crosstalk ratio}=\frac{8\times10^{12}}{\omega KZ}$$

when K is in micromicrofarads and $\omega=2\pi f$, f being in cycles per second and Z the characteristic impedance of the pairs. Converted into decibels (db)

$$\text{crosstalk ratio}=20\log_{10}\frac{8\times10^{12}}{\omega KZ}\text{ db}$$

A similar line of reasoning shows that the total magnetic induction is the resultant of the four mutual inductance couplings shown in Fig. 8(E). If M is the effective resultant magnetic induction, Fig. 8(F)

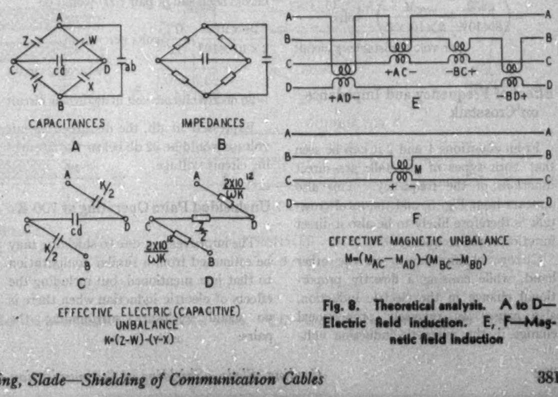


Fig. 8. Theoretical analysis. A to D—Electric field induction. E, F—Magnetic field induction

$$M=(M_{ac}-M_{ad})-(M_{bc}-M_{bd})$$

where M_{ac}, M_{ad}, etc., are the mutual inductances between corresponding wires of the circuits. In this case, the magnetic induction between circuit AB and circuit CD becomes

$$\text{crosstalk ratio}=20\log_{10}\frac{2\times10^6\times Z}{\omega M}$$

where M is in microhenrys.

Total Crosstalk

If the crosstalk ratio is inverted, the fraction (voltage in disturbed circuit)/(voltage in disturbing circuit) will give the voltage in the disturbed circuit for 1 volt applied to the disturbing circuit. The voltages for electric induction and magnetic induction may then be added to give total induction for 1 volt in the disturbing circuit.

For electric induction

$$\text{voltage in disturbed circuit for 1 volt in disturbing circuit}=\frac{\omega KZ}{8\times10^{12}} \quad (1)$$

For magnetic induction

$$\text{voltage in disturbed circuit for one volt in disturbing circuit}=\frac{\omega M}{2\times10^6\times Z} \quad (2)$$

In an electrically short line, where the phase change along the length is small and eddy currents in shields or metal sheath do not add a resistive power component to the mutual inductance, the electrically and magnetically induced voltages are both in phase quadrature with the disturbing voltage. It has been shown that these two voltages add at the near end of the cable, with reasonable accuracy at least up to above 1.0 mc and down to frequencies at which the impedance of the circuit is essentially resistive.

$$\text{total crosstalk voltage at near end}=\left(\frac{\omega KZ}{8\times10^{12}}+\frac{\omega M}{2\times10^6\times Z}\right)\text{ volts per volt in disturbing circuit}$$

Effect of Frequency and Impedance on Crosstalk

From equations 1 and 2 it can be seen that both types of crosstalk are direct functions of the frequency. This also appears from Fig. 5, and the total crosstalk is therefore likely to be also a direct function of the frequency.

Change of impedance, on the other hand, while causing a directly proportional change in the electric induction, also causes an inversely proportional change in the magnetic induction voltage. If the impedance of the circuit is doubled, the current is halved and thus the term ωMI, the voltage due to magnetic induction, is also cut in half. The way in which a change of impedance is reflected in the total crosstalk will therefore depend on which factor is greater, electric or magnetic. If they are equal there will be practically no change when the impedance is varied.

Analysis of Crosstalk in Shielded Pair Cables

The purpose of the rather detailed analysis of crosstalk is to determine the effect of shielding on crosstalk between pairs.

As an example of pairs with copper braid shields operating at 100 kc, consider the case of cable pairs shielded with a single 0.0063-inch copper wire braid applied with 100 per cent coverage, operating at a frequency of 100 kc. The electric field induction is zero if this braid is kept at ground potential. With the use of data from the appropriate curve of Fig. 3, the magnetic field is reduced by the shielding in a ratio of 20 to 120 or 1/6, and the ratio is 1/2×1/6, or 1/12 if both pairs are shielded.

$$\text{total crosstalk voltage}\left(=\frac{\omega}{2\times10^6\times Z}\right)\times\left(\frac{M}{12}\right)\times V$$

where V is the voltage applied to the disturbing circuit and M is the resultant unbalance mutual inductance for this particular cable when unshielded, this being the unbalance due to unavoidable irregularities in manufacture.

In a 1,000-foot length of cable a typical value for M on unshielded cable would be 0.1 microhenry, and a characteristic impedance of 100 ohms. In this case, the total crosstalk voltage (voltage appearing across near end of pair CD) would be

$$\frac{2\pi\times10^5}{2\times10^6\times10^2}\times\frac{0.1}{12}\text{ volts per volt in disturbing circuit}$$

$$=25\text{ microvolts per volt in disturbing circuit}$$

Expressed in db, the disturbed circuit voltage would be 92 db below the disturbing circuit voltage.

Unshielded Pairs Operating at 100 Kc

The improvement due to shielding may be estimated from a similar computation to that just mentioned, but including the effects of electric induction when there is no metal enclosure surrounding the pairs.

$$\text{crosstalk voltage due to electric induction}=\frac{\omega KZ}{8\times10^{12}}\text{ volts per volt in disturbing circuit}$$

A typical value of K for a 1,000-foot length of cable would be 20 micromicrofarads.

$$\text{electric induction}=\frac{2\pi\times10^5\times20\times10^2}{8\times10^{12}}\text{ volts}$$

$$=150\text{ microvolts per volt in disturbing circuit}$$

$$\text{magnetic induction}=\frac{\omega M}{2\times10^6\times Z}$$

$$=\frac{2\pi\times10^5\times0.1}{2\times10^6\times10^2}$$

$$=300\text{ microvolts per volt in disturbing circuit}$$

$$\text{total crosstalk, near end}=150+300$$

$$=450\text{ microvolts per volt in disturbing circuit}$$

$$\frac{\text{disturbing circuit voltage}}{\text{disturbed circuit voltage}}=\frac{10^6}{450}=69\text{ db}$$

The application of copper braid shielding in the example just analyzed would therefore reduce the disturbed voltage from 450 microvolts per volt to 25 microvolts per volt. While the estimated unbalances are approximate, the example illustrates how the advantages accruing from the application of a shield may be evaluated. In terms of db the advantage is 92−69=23 db.

Comparison of Short and Long Lines

The foregoing analysis of the nature and amount of crosstalk is confined to short lines and applies only in a fundamental way to lines of many miles in length. The actual inductive interference for each foot of a long length is subject to the same methods of calculation as already described, but in a long line the total of all these units of crosstalk is not their arithmetic sum. On its way to the line terminations the crosstalk currents are attenuated and subject to a change in phase. The disturbing circuit currents and voltages are similarly attenuated.

If a short line is terminated in a 100-ohm noninductive load, there will be no change in impedance of the circuit as the frequency is varied, for looking into the line from one end we shall always see an impedance of nearly 100 ohms, even at very low frequencies. In this case the electric and magnetic induction remain in the same proportion to each other at high and low frequencies. An average

Vol.
Pt. 1
195

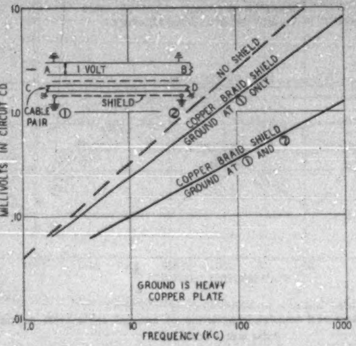


Fig. 9. Total induction tests showing effect of grounds applied to shield over cable pair

pair such as used in multiconductor cable has an attenuation of only 2 or 3 db per mile in the audio range so that the basic equations and calculations are valid in many cases on lines up to several thousand feet in length. In longer lines it has been shown that the impedance seen from one end increases several times at the lower audio-frequencies.

Other Methods of Reducing Crosstalk

While this paper deals primarily with reduction of crosstalk by shielding, other well-known devices for reducing crosstalk should be mentioned. Improvement in uniformity of construction has done much to reduce crosstalk in spiral-four cable. The use of different lays in adjacent pairs has been responsible for improvement in the same direction.

Physical separation of pairs either in the same or different cables may be resorted to when necessary. Crosstalk may be "balanced" out to a large extent by connecting together lengths of cable in such a way that the crosstalk voltages are of opposite polarity and tend to neutralize each other. Also networks may be connected to the ends of the cable to improve the balance.

Of all these methods, shielding is one of the most satisfactory. It must be applied with regard to the nature of the crosstalk, whether it is electric or magnetic. As has been shown, frequency and impedance of the circuit are also determining factors in deciding what type of shielding will be selected. It is certain that there are many cases where expensive copper braids are applied although equal shielding efficiency could be obtained with metallized paper or semiconducting coverings. There are also places where metallized paper is used without regard to the fact that it may be contributing very little to screen out induction where the induction is mainly magnetic.

Grounding of Shield on Shielded Cable

If the parallel wires of circuit *CD* are enclosed in a copper braid shield grounded at both ends we have seen that there is a reduction in voltage measured at the ends of the circuit at *C* or at *D*. This reduction is due to 1. the shielding out of the electric field, and 2. the reduction of the magnetic field because of eddy currents in the braid. Fig. 9 shows two curves for copper braid, the lower one for the braid grounded at both ends, and the one above it for a ground at only one end. When the braid is grounded at one end only, point 1 in the circuit of Fig. 9, there is a rise of voltage along the braid which amounts in the case of the 10-foot test circuit to 14 millivolts per volt applied to the *AB* circuit at 1 mc. When the disturbed circuit *CD* also has one side grounded, there will be a capacity current flowing from the shield to the ungrounded side and the crosstalk will be increased.

The lower curve for the ground-at-both-ends condition, represents essentially the case where electric induction is eliminated and only the magnetic induction remains, which is reduced by the presence of the magnetic shielding effect of the metal braid. The need for frequent grounding becomes less important at the audio-frequencies.

Shielding of Coaxial Cables

The problems involved in shielding coaxial cables are somewhat different from those encountered in shielding pairs. In this case, we are concerned with the current flowing in the disturbing circuit. If two coaxial cables run side by side, as in

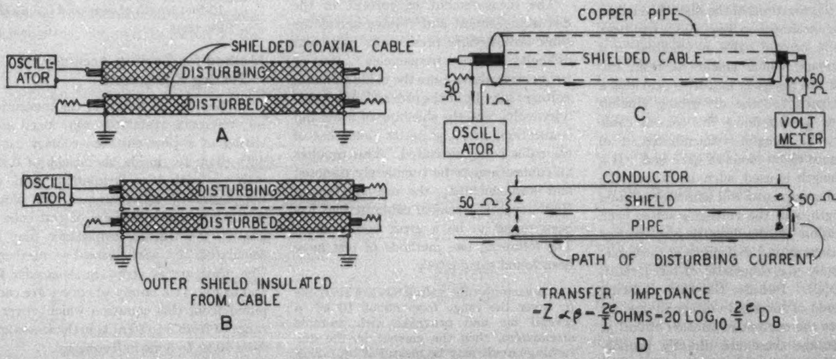


Fig. 10. A,B—Circuits showing inductive interference or crosstalk between coaxial cables. C,D—Circuit for measuring transfer impedance of coaxial cables

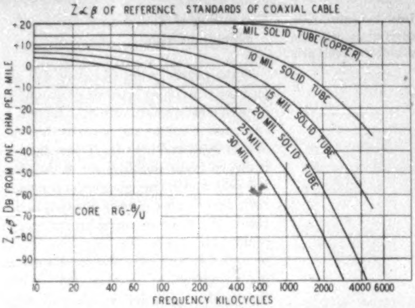


Fig. 11. Calculated transfer impedance of coaxial cable with copper tube outer conductor

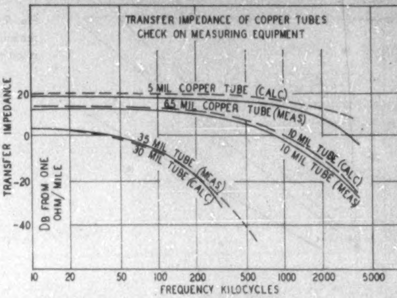


Fig. 12. Measured transfer impedance of coaxial cable with copper tube outer conductor

Fig. 10(A), the sheaths will usually be connected together, and probably grounded at each end. The current in the sheath of the disturbing circuit will therefore have a parallel path through the disturbed cable.

A practice becoming more common is to use "triaxial" cables, with an outside sheath which is grounded but insulated from the inside sheath, the inside sheath carrying the main current of the coaxial cable. Fig. 10(B) shows that there is no physical connection between the disturbing and disturbed cable, except that the isolated outer sheath of the disturbed cable is grounded and the disturbing cable may also be grounded.

There are several other ways in which interaction may occur between two coaxial cables, but they all result in the application of voltage, either directly or by induction, across the outer braid of the disturbed cable. To determine the voltage that a given disturbance will produce at the terminations of the disturbed cable, the circuit shown in Fig. 10(C) has been used. A coaxial cable is enclosed in a copper pipe which shields it from unwanted interference, and also provides a return path for the disturbing current which is passed down the coaxial cable shield. The coaxial is terminated in an impedance of 50 ohms at each end. If a short length is used, such as 10 feet, the disturbing current I will be uniform along the length, and the voltage e across each termination will be directly proportional to the current and the length of the cable. Fig. 10(D) is a schematic of test circuit, Fig. 10(C). Because the ratio is being considered of the total voltage in one circuit α to the current in another circuit β, and because these are directly proportional, the relationship between these two circuits may be defined as a transfer impedance

$$Z_{\alpha\beta}=\frac{2e}{I}\text{ ohms}$$

$$\text{or } Z_{\alpha\beta}=20\log_{10}\frac{2e}{I}\text{ db}$$

where $2e$ = total voltage in the disturbed circuit. For convenience in applying the results to system calculations, the values are converted to volts per mile for 1 ampere of disturbing current flowing in the shield. The test results given which follow are plotted as decibels above or below 1 ohm per mile.

An analysis of surface transfer impedance based on the electromagnetic wave theory is given by Schelkunoff.[4] Extracts are given in Appendix I.

Procedure for Making Tests of Shielding Efficiency of Coaxial Cable

The measurement of current in the disturbing circuit and voltage across the cable terminations presents no particular difficulty at low frequencies. But at frequencies above 1 mc the cable terminal voltages may be of the order of only 1 or 2 microvolts, and the shielding of leads and connections must be better than that of the cable being measured. This requires all connections to be completely enclosed and well shielded; the use of General Radio type-874 coaxial cable fittings has been found to be a great convenience. The following two methods of test have been found satisfactory.

1. If vacuum-tube voltmeters are available to cover the range from about 10 kc to several mc and provided with suitable attenuators, then the current in the disturbing circuit may be measured by taking the drop across a 1-ohm noninductive resistor in the braid circuit. The drop across the 50 ohm termination is also measured directly and this drop, multiplied by 2, calculated for a 1-ampere disturbing current gives the $Z_{\alpha\beta}$ or transfer impedance for the length of cable being tested.

2. Alternatively the voltage across the 1-ohm resistor may be attenuated until it is equal to the drop across the 50-ohm resistor. Radio receivers may be used for checking the equality of the 50-ohm signal and the attenuated signal across the 1-ohm resistor, and in this case there is no need for the radio receiver to be calibrated.

In this test the attenuator reads directly the ratio e/I in db, this must be multiplied by 2 and have the length correction applied.

$$Z_{\alpha\beta}=\frac{e}{I}\times\frac{2\times5{,}280}{L}\text{ ohms per mile}$$

where e and I are as in the foregoing and L is the length in feet of cable under test.

$$Z_{\alpha\beta}=\left(30\log_{10}\frac{e}{I}+20\log_{10}\frac{2\times5{,}280}{L}\right)\text{ db}$$

$=(-\text{attenuator reading}+60\text{ db})$ for the 10-foot length of pipe used for making the tests

Method for Testing Accuracy of Testing Equipment

To test the accuracy of the measuring set, reference standards were used consisting of copper shielded coaxial cable, with diameter inside the shield of 0.260 inch, polyethylene insulation, and an impedance of approximately 50 ohms. When the wall thickness of the tube is known, the transfer impedance may be calculated by Schelkunoff's method.[4] The equation is given in Appendix II, and in Fig. 11 a family of curves are computed from this equation which cover a range of from 5 to 30 mils in thickness and from 10 kc to 5 mc in frequency.

Copper tubes of 6.5-mil wall, 10-mil wall, and 35-mil wall were available for

Vol. 7
Pt. 1
1955

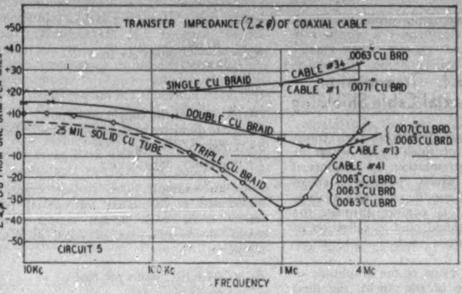


Fig. 13. **Measured transfer impedance of coaxial cable with copper braid shields**

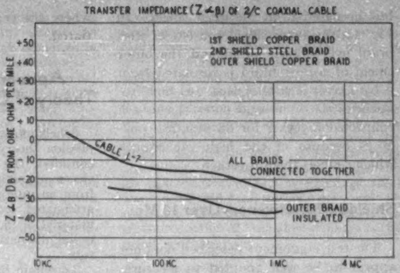


Fig. 15. **Measured transfer impedance of triaxial cable. Outer braids insulated from each other**

test, and the curves of Fig. 12 indicate the margin between calculated and measured values. The agreement is good considering that the conductivity of the copper tubing was assumed to be 100 per cent.

Shielding Effectiveness of Metal Braids and Tapes on Coaxial Cable

The shielding efficiency of flexible metal braids or tapes applied to coaxial cable cannot be determined by calculation because of the lack of continuity of the covering, radiation through spaces between wires or tapes, uncertainty of contact, resistance of overlapping wires of the braid, and several other factors. It is therefore necessary to make actual tests with apparatus as described in the foregoing to evaluate the efficiency of other than the simplest tubular metallic coverings, in screening out electromagnetic radiation. These tests have also to cover the frequency range of the disturbing field under consideration, for two shields are compared, one may be better at low frequencies and the other superior at high frequencies.

Comparison of Single, Double, and Triple Braid Shielding

Fig. 13 shows the improvement obtained by the addition of a second and a third braid over a single-braided coaxial cable. At 1 mc the triple braid is 56 db better than the single braid, so that the addition of the two extra braids has reduced the induced voltage in the coaxial cable circuit almost to one-thousandth of its original value. But at 4 mc, the radiation through the braid spaces has evidently increased and is in phase with the attenuated magnetic field which has penetrated the material of the shield itself. This aspect of the subject is discussed briefly in Appendix I.

Combination Shields of Copper and Steel Braids

Reference is made in Appendix I to the fact that when the electromagnetic field passes from one medium to another of different impedance, a part of the field is reflected, just as it is in a transmission line, the reflection coefficient being proportional to $\sqrt{ug}$. It was found that a considerable gain in shielding could be obtained by the use of combination steel and copper braids. In Fig. 14 the copper-steel-copper braid of cable 26 is a great improvement over the triple copper braid. Even at 10 kc it has a 20-db advantage due almost certainly to the presence of the steel, and its shielding efficiency is high to a frequency above 4 mc.

Substitution of Steel Tapes and Aluminum Braids in Place of Copper and Steel Wire Braids

Braiding is basically a slow operation and it was proposed that wrapped steel tapes could be applied more economically and perhaps more effectively than a steel braid. Also aluminum braid was tried as a substitute for copper braid. The combination copper braid-steel tapes-aluminum braid, when applied with the two steel tapes wrapped in opposite directions approached the limit of the testing equipment above 1 mc, evidently because the radiation through the minute braid spaces was almost completely checked by the steel tape barrier.

Triple Shielded Cable

Where the situation permits, the cable with two shields insulated from each other as in Fig. 15 offers greatly improved shielding, especially at low frequencies. If the outer shield is grounded, the inner shield can be grounded only at one end. Some of the advantage of isolating the outer shield is no doubt lost on long lengths and at high frequencies. In Fig. 10(B) the outer shield is shown grounded

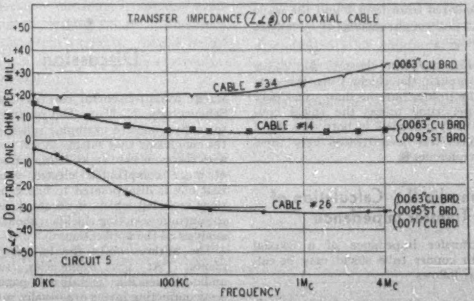


Fig. 14. **Measured transfer impedance of coaxial cable with copper and steel braids**

but insulated from the inner shield. In Fig. 15, the inner shield consists of steel braid over copper braid and the outer shield of copper braid is insulated from it by a layer of rubber tape. At high frequencies the inner and outer shields are connected together for all practical purposes because of the high capacitance which couples them together.

Shielding Effectiveness Over 10 Mc

The test equipment described does not appear to be well adapted to measurements of over 10 mc. If a shielded cable is exposed to radiation from an antenna, the conditions of test must be rigidly prescribed. Distance of antenna from cable, angle between antenna and cable, height above ground, and several other factors all influence results to a marked extent.

Tests have been reported by Shive[5] and others dealing with the value of shields placed between an energized primary coil and a secondary coil receiving the electromagnetic energy which extends into the range above 100 mc. Work is being undertaken to evaluate cable shields exposed to antenna radiation, in the same range of frequency.

In conclusion, an attempt has been made to show that to operate efficiently, shields should be chosen with the same care which is expended in the selection of more complicated electric equipment.

At low frequencies, such as audio-frequencies, metallized paper and semiconducting layers are usually sufficient, for while entirely ineffective against magnetic induction, they eliminate electric induction, and the magnetic component is often unimportant just because the frequency is low. The use of copper braids at these frequencies is usually economically unwarranted.

In the carrier range steel tapes and braids, or a lead sheath or steel pipe will considerably reduce the induction. The current in the circuit should be kept as low as possible by raising the value of the terminating impedance or keeping the cable impedance high. Transposing or physically separating the circuits are also important in reducing induction.

At high-frequencies coaxial cable circuits are more commonly used, and where the best possible shielding is essential, heavy copper tubing of 30 mils or more in thickness or the multiple copper-steel-copper braid provides a high degree of protection.

Much of the work described here was carried out under the auspices of the Signal Corps in co-operation with engineers of the Coles Signal Laboratory, where the coaxial cable project was initiated.

Appendix I. Transmission Theory of Coaxial Cable Shielding

Schelkunoff[4] has developed the "transmission theory" of shielding in which he considers the attenuation of cylindrical electromagnetic waves passing outward from an internal conductor. The wave is attenuated by: 1. passing through the material of the shield; and 2. reflection at radial discontinuities in the shield structure.

To obtain the value of the amplitude of the wave outside of the shield, the final value of the attenuated wave from passing through and from reflection must be added to the radiation loss through any interstices that may occur in the shield. These values may not be in phase, so that it is possible for radiation loss through interstices actually to neutralize the field which passes through the body of the shield, at certain frequencies.

The reflection loss at shield boundaries is frequently greater than the loss due to attenuation, which emphasizes the importance of using shields of different materials. The reflection loss is proportional to the value K', which is the impedance ratio for the two adjacent materials. The impedance ratio for two metals is

$$K' = \frac{u_1 g_2}{u_2 g_1}$$

where u_1 and u_2 are the permeabilities and g_1, g_2 are the conductivities of the metals. This ratio is independent of the frequency.

For the attenuation loss through the metal, $\alpha = \sqrt{\pi u f g}$, which gives a value increasing as the square root of the frequency. The frequency in cycles if f, u and g are as in the foregoing, and α is the attentuation per unit thickness of the metal.

The total value of the field immediately outside the cable will therefore be computed from.

1. The total attenuation of the incident wave on the inside of the shield by reflection and by attenuation through the thickness of the metal.
2. Phasorial addition of the attenuated wave calculated from item 1 and the wave which passes through openings or interstices in the shielding.

The same procedure is used for waves passing through the shield from the outside to the inside, but the final value may not be the same unless the shield construction is symmetrical in both directions. This subject has been treated very thoroughly in reference 6.

Appendix II. Calculation of Transfer Impedance

The transfer impedance of a coaxial cable with copper tube shield may be calculated as follows

$$Z_{\alpha\beta} = R_{dc} \frac{u}{\sqrt{\cosh u - \cos u}}$$

where

$$R_{dc} = \frac{10.58}{4t(d+t)} \text{ ohms per foot}$$

where

t = thickness of the copper tube, mils
d = inside diameter of the copper tube, mils
$u = 0.7686t\sqrt{F}$ radians
F = frequency, megacycles

As an example, the surface transfer impedance is calculated for coaxial reference cable with 30-mil copper tube shield—inside diameter of 0.260 inch at 2.7 mc.

$$R_{dc} = 3.04 \times 10^{-4} \text{ ohms per foot}$$
$$= 1.61 \text{ ohms per foot}$$
$$u = (0.7686)(30)\sqrt{2.7} = 37.9 \text{ radians}$$
$$\cosh u = {}^1/_2(282 \times 10^{14}) = 141 \times 10^{14}$$

Since cos u can never be larger than 1 it can be neglected in the equation for surface transfer impedance when cosh u is very large.

$$Z_{\alpha\beta} = (1.61)\frac{37.9}{\sqrt{141 \times 10^{14}}} = 5.11 \times 10^{-7} \text{ ohms per mile}$$

Converting $Z_{\alpha\beta} = 20 \log_{10} 5.11 \times 10^{-7} =$ 126.2 db from 1 ohm per mile at 2.7 mc.

References

1. RADIO ENGINEERS' HANDBOOK (book), **Frederick E. Terman.** McGraw-Hill Book Company, Inc., New York, N. Y., 1943.

2. HOW CROSSTALK AND NOISE AFFECT CAPACITANCE UNBALANCE REQUIREMENTS, PARTICULARLY IN SPIRAL-FOUR CABLES, **A. G. Chapman.** Signal Corps Symposium, Ft. Monmouth, N. J., 1952.

3. CROSSTALK, **H.C.S. Hays, R.A. Seymour, P.R. Bray.** British Post Office Engineering Department, London, England, 1938.

4. ELECTROMAGNETIC WAVES (book), **Sergi A. Schelkunoff.** D. Van Nostrand Company, Inc., New York, N.Y., 1948.

5. EFFECTIVENESS OF CONDUIT AS R.F. SHIELDING, **Scott L. Shive.** *Electronics*, New York, N. Y., vol. 19, Feb. 1946, pp. 160–66.

6. STUDY OF TUBULAR SHIELDING AND DEVELOPMENT OF FLEXIBLE SHIELDING CONDUIT, **A. R. Anderson.** *Final Report Contract No. a-s-8098*, United States Bureau of Aeronautics, Washington, D. C.

◆

Discussion

H. R. Meahl (General Electric Company, Schenectady, N. Y.): The paper contains so many concrete examples which support the conclusion that shields should be chosen with the same care expended in the selection of more complicated electric equipment that one is disappointed to find generalizations included which, if accepted, lead one to continue selecting shields without careful analysis of the application.

The authors make the following statements: "At low frequencies, such as audio-frequencies, metallized paper and semiconducting layers are usually sufficient, for while entirely ineffective against magnetic induction, they eliminate electric induction, and the magnetic component is often unimportant just because the frequency is low. The use of copper braids at these frequencies is usually economically unwarranted." These statements contradict the discussion in the section "Magnetic Field Shielding," the data of Fig. 3, and general experience.

The authors gave examples of the use of steel braid, both alone and combined with copper tape, but did not point out that dissipative shielding is a class of shielding which is also worthy of serious study. Steel wool and a wide variety of compounds containing powdered iron and carbon are quite effective in converting electromagnetic energy into heat therefore preventing it from re-entering a communications circuit, as energy which escapes through openings in supposedly complete shields so often does.

It is hoped that readers will be led to act upon the main thesis that shielding is not a simple problem, rather than lulled into complacency by the several statements which tend to give that impression.

F. H. Gooding and H. B. Slade: We believe that the use of copper braid is not usually justified for shielding purposes in the audio range because the required shielding against electric induction can be more economically effected by the use of metallized paper or semiconducting tape with a drain wire. Occasionally metal braid is used for mechanical purposes, or where very positive ground connections are desirable.

We regret having given the impression that the selection of shields is a simple matter; and agree with Mr. Meahl that in a short paper it is not possible to cover the subject comprehensively. Dissipative and absorptive shielding obtained through the use of thick semiconducting coverings, steel wool, powdered iron, carbon, etc., may be especially effective at high frequencies, where reduction of "opening" leakage is the prime consideration.

Progress in the Evaluation of Solid-Core High-Voltage Bushings

H. H. BRUSTLE
ASSOCIATE MEMBER AIEE

D. L. JOHNSTON
MEMBER AIEE

A. L. SCHEIDELER
ASSOCIATE MEMBER AIEE

l, 74
1
55

A HIGH-VOLTAGE bushing which successfully sustains repeated application of 60-cycle and impulse test voltages may not be adequate for prolonged operation at rated line-to-ground voltage. Methods to evaluate bushings at voltages slightly in excess of operating voltages have therefore been devised. The techniques developed have been used to evaluate the life characteristics of the solid compound core type of bushing and to provide a routine quality control of bushing production.

Most responsible writers on resin-bonded paper dielectrics recognize that the finished product may have air inclusions. If voids are located in areas of high electric stress, harmful levels of ionization may occur at operating voltage. These gaseous discharges, generally known as corona, may contribute to radio and television interference and may eventually cause bushing failure. Therefore, the corona characteristics should be determined at operating voltage as a routine factory procedure.

Three methods of measuring corona within the single equalizer type of bushing have been evaluated. These tests include measurement of radio noise in microvolts using standard circuits,[1] measurement of corona intensity in micromicrocoulombs using the resonant circuit technique,[2] and power factor (pf) measurements as a function of voltage.[3]

Single-Equalizer Design

Solid compound bushing cores employing a single equalizer near the ground plane have been in service since 1939. From 1932 to 1939 the design was identical except that external bell-shaped metallic ground shields were used in place of the equalizer. This single-equalizer design, illustrated in Fig. 1, has proved practical and has achieved a very satisfactory service record. The small percentage of service troubles traceable to void ionization have been eliminated by improvements in the methods of initial evaluation.

As the ratio of internal to external conductor radii increases, the need to employ a multiplicity of conducting layers, as used in the designs internationally known as condenser bushings (Fig. 2), to equalize radial gradients becomes apparent. In fact, where an oil-permeable insulation structure is used, many conducting planes may be advocated. From a practical manufacturing standpoint, there are only two logical design schemes when synthetic resin bonded material is used. These require either as few conducting planes as possible, as illustrated in Fig. 1, or so many that corona is prevented even when voids are present. Extensive service experience clearly indicates that satisfactory axial and radial gradients may be obtained in ratings through 69 kv with the single equalizer design.

Corona Intensity Measurements with Use of Resonant-Circuit Techniques

The detection of ionization has received a great deal of attention over the past few years as has the phenomenon involved during the ionization of air voids. Not until recently, however, was a practical method evolved which would allow the measured quantity to be related to the amount of charge which flows during the ionization process.[3]

Briefly, the method involves placing an inductance coil in series with the test piece and ground; see Fig. 3. The high-voltage connections as well as the testing transformer must be corona-free. When corona occurs in the voids in the test piece it disturbs the energy balance of the circuit. The redistribution of charge which takes place causes a small voltage to appear across the terminals of the measuring inductance. The signal is then amplified and displayed on an oscilloscope. By proper calibration, the voltage of the

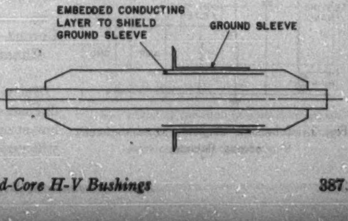


Fig. 1 (right). Cross section, single equalizer bushing core

Paper 55-176, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing November 29, 1954.

H. H. Brustle, D. L. Johnston, and A. L. Scheideler are with the General Electric Company, Pittsfield, Mass.

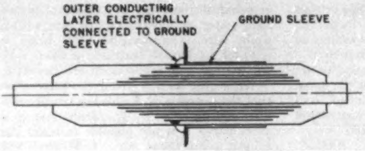


Fig. 2 (left). Cross section, condenser bushing core

actual corona pulse can be measured and related to the charge transferred in the discharge process.

Radio Noise Measurements

Standard conducted radio noise measurements[1] provide a simple technique for measuring corona severity. Although the sensitivity of the noise meter method is a function of sample capacitance, this is not a particular handicap with the low-capacitance, single-equalizer type of core. While the resonant-circuit corona detector provides a more definitive tool for developmental evaluation, the inherent simplicity of the radio noise test makes it a valuable and practical tool for production testing. Also, since many utilities have means of obtaining radio noise data such results can easily be compared with manufacturers' initial results. The comparison of such results provide valuable life test data as well as a possible basis for determining continued service life of the bushing.

PF Versus Voltage Measurements

It has been proposed[3,4] that conventional Schering bridge pf measurements may serve as an index of corona within a piece of electric apparatus. Briefly, the proponents of this method maintain that a sharp upward bend in the pf versus voltage curve is indicative of damaging corona levels in the dielectric.

To evaluate the meaning of such tests it is necessary to consider the significance of the pf test itself. The pf at any voltage and frequency is equal to the ratio, watts dissipated/volt amperes, which is the sine of the loss angle. Thus, any property of the material which contributes to energy loss as voltage is applied to the core will, in general, contribute to the pf. For comparatively small angles, the pf is approximately equal to the tangent of the loss angle. This tangent, which is known as the dissipation factor, may be expressed as $G/\omega C$ where G is the equivalent parallel conductance, $\omega = 2\pi \times$frequency, and C is the parallel capacitance of the circuit. Each of these factors might conceivably be affected by voltage in such a way as to affect the slope of the pf versus voltage curve. In succeeding sections each will be considered in the light of available information.

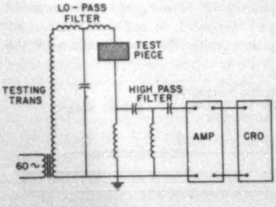


Fig. 3. Schematic diagram, resonant-circuit corona detector

Characteristics of Single-Equalizer Core Structure

To obtain directly comparable information on the pf versus voltage, corona, and radio noise-measuring techniques, comprehensive tests have been performed on a core structure of the single-equalizer type.

Pf, capacitance, corona and radio noise tests were made at each voltage level. A switching arrangement was used to connect the test piece to the Schering bridge for pf and capacitance measurements, to the corona detector for corona intensities, or to the Ferris *32B* radio noise meter for radio influence voltage levels.

The procedure was used at temperatures ranging from 5 to 80 degrees centigrade (C) in order to establish the effect of temperature on each test and to assess the repeatability of such measurements where ambient is not controlled. Room temperature measurements of corona and radio noise characteristics have also been made on cores at voltages as high as the high-potential test voltage.

Corona and Radio Noise Characteristics

The measured intensities of the corona discharge, as illustrated in Fig. 4, are very low at operating voltage and, in general, still very low at nominal voltage. The effects of temperature on the corona-starting voltage are negligible in the range considered.

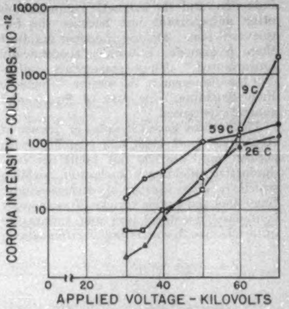


Fig. 4. Corona versus voltage at various temperatures; 69-kv bushing; operating voltage 39.8 kv

The corona characteristics of the single-equalizer type of bushing structure described in this paper differ considerably from those of liquid-permeated structures. It has been shown[2] that for the latter, ionization characteristics versus voltage can reveal incipient failure of the insulation. Typical information of this nature obtained on capacitors show that such structures possess a so-called break-point in the corona intensity curve. Continued operation at voltages over this point is damaging in a relatively short time. This break-point type of corona intensity characteristic has been observed in other liquid-solid dielectric systems such as restricted barrier-oil gaps of uniform and nonuniform fields. There again, as is the case with capacitors for safe design, the high-potential test stress should be substantially below the stress at which this break-point in corona intensity occurs.

Corona levels of the single-equalizer type of core structure, in contrast with

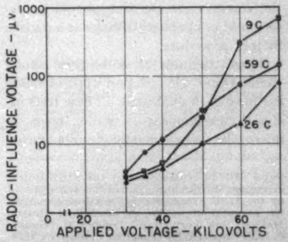


Fig. 5. Radio noise versus voltage at various temperatures; 69-kv bushing; operating voltage 39.8 kv

Vol
Pt.
195

l, 74
1
55

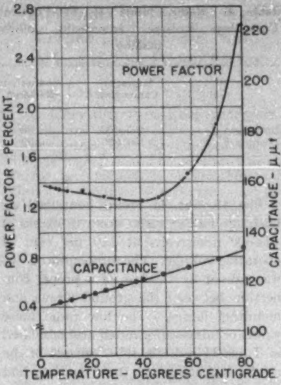


Fig. 6. Capacitance and pf versus temperature; 10 kv

those for liquid-filled systems, vary gradually with increasing voltage. The shape of the curve, which, is greatly influenced by void dimensions and position, is typical of the restricted ionization characteristic associated with voids in a solid structure. It represents the ionization level of the weakest link in the dielectric at any given voltage. This characteristic curve is quite reproducible in the normal product, indicating that no damage has occurred. Referring again to Fig. 4, it will be noted that ionization begins near operating voltage. As voltage is raised, voids located throughout the structure are ionized. Actually, the pattern observed on the oscilloscope usually gives a crude indication of the number of air voids present.

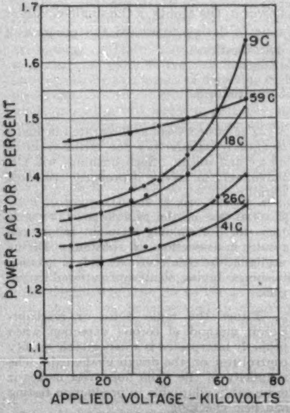


Fig. 7. Pf versus voltage at various temperatures; 69-kv bushing; operating voltage 39.8 kv

The radio noise characteristics of this typical core are shown in Fig. 5. A very close similarity is exhibited by these data and the corona intensities of Fig. 4. It is evident that they are providing a measure of the same characteristics. Because of these similarities, and due to the facility with which the radio noise measurement may be effected, this test is now in use as a routine quality control for solid compound bushings.

It is apparent that some relation must exist between the corona intensity level and the service life of the insulation. Although the relationship is not as sharply defined as it is for the liquid-filled structures, significant conclusions may be drawn from accelerated life tests, field experience, and standard production testing.

Capacitance and PF Characteristics

The capacitance of the bushing core is independent of voltage between zero and operating levels, throughout the test range from 5 C to 80 C. It may be concluded from this that nonlinear changes in capacitance which would influence the shape of the pf voltage curve are not present. The change in capacitance with temperature, however, is quite significant. As may be seen in Fig. 6 the capacitance of the core is a direct function of the absolute temperature. Such capacitance changes are responsible in large part for a decrease in pf with temperature in the region, 5 C to 45 C (Fig. 6). Above 45 C there is a substantial increase in pf level with temperature. This is apparently due to an increase in ionic conduction which is influenced by the decrease in "viscosity" with rising temperature of the compound resin.[1] Such temperature effects make it essential that comparative pf versus voltage measurements on solid bushings be made at carefully controlled temperatures.

The relationships between pf and voltage are illustrated in Fig. 7. It is obvious that the shape of the pf versus voltage curve is greatly influenced by the test temperature. The position, and even the existence, of a bend in the curve are dependent upon temperature level. If the shape of the curve were being influenced by changes in corona, such effects should be reflected in the measurements of corona and radio noise as a function of temperature. It may be seen from Figs. 4 and 7 that such is not the case. Measured corona levels are higher at 59 than at 26 C, while the pf bend is pronounced at 26 and nonexistent at 59 C.

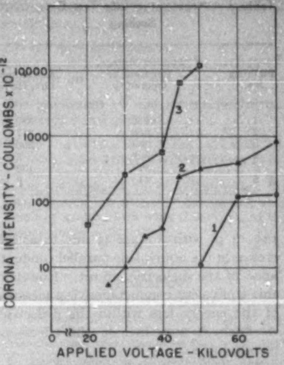


Fig. 8. Effect of variations in core structure on corona versus voltage characteristics; 69-kv bushings; operating voltage 39.8 kv

Contribution of Corona Losses to Dielectric Losses as Measured by Pf

Since the effects of voltage on capacitance were negligible, it follows that any

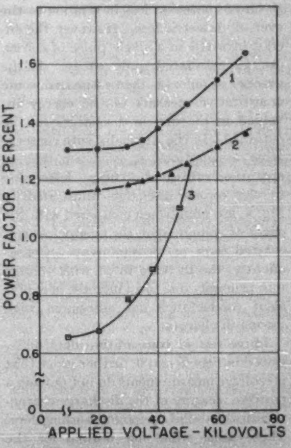


Fig. 9. Effect of variation in core structure on pf versus voltage characteristics; data from same cores as those of Fig. 8

Table I. Accelerated Life Test Data, 69-Kv Bushing

Bushing No.	Applied Voltage, Per Cent of Operating	Accelerated Life, Hours
1	250	330
2	300	317
3	338	200
4	338	5
5	375	130
6	440	6

rise in pf with voltage is due to an increase in the equivalent parallel conductance of the dielectric circuit. However, this equivalent conductance, is a measure of the energy loss within the dielectric. As such, it may include the effects of corona losses.

There is considerable information in the literature regarding the change of conductivity with voltage.[6] Investigators who have studied the phenomenon are, however, in wide disagreement as to the mechanisms involved. Consequently, there is no generally accepted way available to separate analytically the conductivity due to some property of the material from that due to corona losses.

But the corona loss component may be separated from the conductivity by measuring or calculating the average energy of the corona discharges. This energy loss may then be subtracted from the over-all energy loss as calculated from pf measurements.

Unfortunately, with the resonant-circuit type of measurement, no indication is given of ionization loss in relation to the over-all dielectric loss. However, the energy involved in a single pulse of corona can be calculated quite easily. If the number of pulses is then estimated, some quantitative measure of the energy involved in the corona discharges can be obtained. In the Appendix with conservative assumptions, it is shown that at operating voltage the energy losses due to the corona discharge, illustrated in Fig. 4, are small when compared with the effect of voltage upon the energy loss calculated from pf measurements. Consequently, the increase in pf with voltage was primarily due to an increase in equivalent conductance not associated with corona discharges.

Three sets of comparative data, illustrated in Figs. 8 and 9, further prove that pf-voltage measurements do not provide a positive measure of the discharge characteristics of a solid compound bushing core. These data, while not normal for the standard product, do show (sample 1, tested at 9 C) that a pf bend may occur while corona is nonexistent and, conversely, that the pf bend may be negligible (sample 2, tested at 25 C) although ionization exists. Data also illustrate that the correlation does exist between pf versus voltage and corona measurements (sample 3, tested at 25 C) when corona is severe.

It is most probable that the increases in pf with voltage are caused by small increases in ionic conductivity. Such changes are common to many dielectrics. It should be remembered that all the pf values being considered are low.

Accelerated Life Testing

A study of results obtained by accelerated life tests on bushings returned from the field indicates that considerable care must be taken in the interpretation of such data. Two groups of 69-kv bushings were recently returned from the field for detailed engineering and laboratory investigation. One group of seven bushings had been in service less than 5 years, while the second lot of nine averaged 18 years in the field.

All bushings in the under 5-year group had radio noise characteristics exhibiting between 2,000 and 4,000 microvolts at 110 per cent of normal operating voltage. Six of these units were placed on life test at voltages significantly greater than operating voltages. In all cases except one, bushing 4 of Table I, the time interval to failure was sufficient to justify optimism about the projected service life of bushings with this much radio noise. The failure path was axial in all cases. One other bushing core of the same category was carefully dissected without the imposition of a voltage greater than it had experienced in service. No evidence of axial trouble was located. Instead, radial deterioration was discovered at the surface of the center conductor. It should be noted that the verified service trouble was radial and the accelerated life test overvoltage caused axial failures.

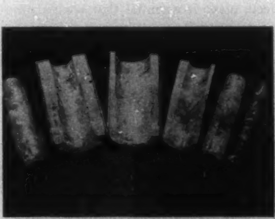


Fig. 10. Beginning of radial failure; 18-year-old 69-kv type-L bushing exhibiting greater than 100,000-microvolt radio noise

Table II. Radio Noise at 110-Per-Cent Operating Voltage, Microvolts, 69-Kv Bushing

	Improper Connection	Connection Repaired
Bushing *A*	>100,000	60
Bushing *B*	>100,000	100

The 18-year-old group contained six bushings having radio noise in excess of 100,000 microvolts at 110 per cent of operating voltage. The radio noise on two was caused by the loss of a proper connection between the ground sleeve and mounting flange. The low radio noise values obtained after repair and tabulated in Table II indicate that, outside of the mechanical deficiency, these bushings had not deteriorated either radially or axially. Three other bushings had deteriorated, and a radial failure path was well started at the conductor but not complete; see Fig. 10. The sixth 18-year-old bushing showed deterioration in the central portion of the core.

The remaining three 18-year-old bushings exhibiting 100, 850 and 1,000 microvolts respectively at 44 kv (National Electrical Manufacturers Association test voltage) were carefully dissected. No evidence of dielectric deterioration could be found in these.

It is apparent that verified service troubles are radial in nature while accelerated life tests at overvoltage cause axial failure. Actually, none of the deterioration mentioned in this section is desirable. These results demonstrate, however, the ability of the single-equalizer type of design to retard the progress of deterioration.

Conclusions

1. The basic design as illustrated in Fig. 1 is entirely satisfactory for bushings rated 69 kv and under. Such bushings will last indefinitely in service if properly evaluated initially.

2. The use of the pf-voltage curve as a basis for determining incipient failure can result in rejection of bushings having negligible internal corona and in passing bushings having significant internal ionization.

3. Either the radio noise or resonant-circuit method of corona detection when properly interpreted is a dependable quality control test on the design evaluated. The simplicity of the radio noise test makes it more easily adaptable to production testing and field checking.

4. It is essential that comparative pf versus voltage measurements be made at carefully controlled temperatures.

Appendix I. Separation of Corona Losses From Those Measured by Power Factor

Corona Losses

Assuming that 1. the dielectric region is initially free of space charge; 2. each pulse occurs substantially at the peak of the voltage wave; 3. the metal electrodes are perfect conductors in comparison with the ionized volume of dielectric; and that 4. ΔE is small with respect to E_p; it has been shown in unpublished data by W. R. Wilson that

$$\Delta W=\frac{CE_p\Delta E}{2} \qquad (1)$$

where

ΔW = energy per pulse, joules
C = capacitance of sample, farads
E_p = peak value of applied voltage, volts
ΔE = instantaneous change of voltage across C caused by the most intense pulse of corona, volts

Also, since

$$Q=C\Delta E \qquad (2)$$

where Q = charge per pulse, equation 1 becomes

$$\Delta W=\frac{E_pQ}{2} \qquad (3)$$

For the single-equalizer bushing core whose characteristics are illustrated in Figs. 5, 6, and 7 at 26 C and 40 kv, rms

$\Delta Q=8$ micromicrocoulombs

$E_p=5.65\times10^4$ volts

So

$$\Delta W=\frac{(8\times10^{-12})}{2}(5.65\times10^4)=0.23\times10^{-6} \text{ joules} \qquad (4)$$

Assume next that there are 1,000 pulses per second, which means approximately 8 pulses per 1/2 cycle. Preliminary measurements indicate this to be a rather high estimate.

$$\text{Ionization loss in watts}=\Delta W\left(\frac{\text{pulses}}{\text{seconds}}\right)$$
$$=(0.23\times10^{-})^{6}(1{,}000)=0.23 \text{ milliwatt} \qquad (5)$$

Losses as Measured by Pf

$$\text{Watts loss}=EI\cos\theta\cong EI\tan\delta \qquad (6)$$

where

E = applied voltage, volts
I = current flow, amperes
δ = loss angle
θ = dielectric phase angle
$\omega=2\pi$ frequency, radians per second
C = capacitance of dielectric structure, farads

Since

$$I\cong E\omega C \qquad (7)$$

$$\text{watts loss}=E^2\omega C\tan\delta$$

For the core at 26 C considered in the foregoing:

$C=117$ micromicrofarads
Tan δ, by extrapolation, at 0 kv = 0.01276
Tan δ at 40 kv = 0.01305

The loss at 40 kv due to changes in tan δ between 0 and 40 kv is

$$\text{Watts loss}=(4\times10^4)^2(377)\ (117\times10^{-12})\text{-}(0.01305-0.01276)=20.5 \text{ milliwatts} \qquad (8)$$

On the basis of these assumptions, the losses due to low intensity corona are very small by comparison with those causing the change in pf with voltage.

References

1. Methods of Measuring Radio Noise, 1940. *NEMA Publication No. 107*, National Electrical Manufacturers Association, New York, N. Y., 1940.

2. Development of Corona Measurements and Their Relation to the Dielectric Strength of Capacitors, **R. J. Hopkins, T. R. Walters, M. E. Scoville.** *AIEE Transactions (Electrical Engineering)*, vol. 70, pt. II, 1951, pp. 1643–51.

3. Bushing Tests, **A. L. Brownlee, W. H. Wickham.** *AIEE Transactions*, vol. 60, Oct. 1941, pp. 940–42.

4. Internal Corona Discharges in Insulated Cables, **E. W. Greenfield.** *AIEE Transactions*, vol. 66, 1947, pp. 591–609.

5. Dielectric Properties of Organic Coatings, **T. W. Dakin.** *Official Digest*, Federation of Paint and Varnish Production Clubs, Philadelphia, Pa., Jan. 1952, pp. 42–56.

6. Dielectric Breakdown of Solids (book), **S. Whitehead.** Oxford University Press, London, England, 1951, pp. 96–103.

◆

Discussion

W. H. Wickham (Commonwealth Edison Company, Chicago, Ill.): The presence of corona discharges in solid-core high-voltage bushings and the resulting deterioration have long been recognized by users concerned with the maintenance of high-voltage bushings. The pf-voltage characteristic was suggested early in the development of bushing testing as a means of detecting such discharges (see ref. 3 of the paper). The need for such a test was not generally accepted. The present recognition of the possibility of bushing deterioration and failure from corona discharges and the use of a corona test for quality control in manufacture and for field checking is a real contribution to improvement of the service reliability of high-voltage bushings of the solid-core type.

While the authors have reached the conclusion that the radio noise and the resonant-circuit methods of corona detection are more dependable than the pf versus voltage test, this is not demonstrated conclusively by the data given in the paper. A different conclusion is reached by the following analysis of the same data.

In the pf versus voltage curve at 9 C shown in Fig. 7 it will be noted that from 10 to 35 kv the curve can be drawn as a straight line. At 35 kv the slope changes. At the higher voltages the difference in pf between this straight line, extrapolated, and the experimental curve should be a measure of the ionization loss. This loss has been calculated for the 9 C curve and the semilogarithmic plot is curve 3 in Fig. 11. The corona-intensity curve in Fig. 4 and the radio-noise curve in Fig. 5 have been plotted (curves 1 and 2) in Fig. 11 for comparison with the ionization-loss curve. It is evident that the three curves have very similar shapes and therefore must be a measure of the same characteristic, i.e., the increase in corona discharge with voltage. The slope of the curve is a measure of the effectiveness of the method in detecting corona discharge. The average slopes of the three curves of Fig. 11 have been tabulated for comparison. The slopes are nearly equal and it follows that the three methods are equally effective. Figs. 12 and 13 show the same comparisons at 26 C and 59 C respectively, and support the foregoing conclusions. Further, a comparison of the slopes at the three temperatures shows that the corona-intensity and radio-noise characteristics have a temperature dependence similar to that of the ionization-loss characteristic.

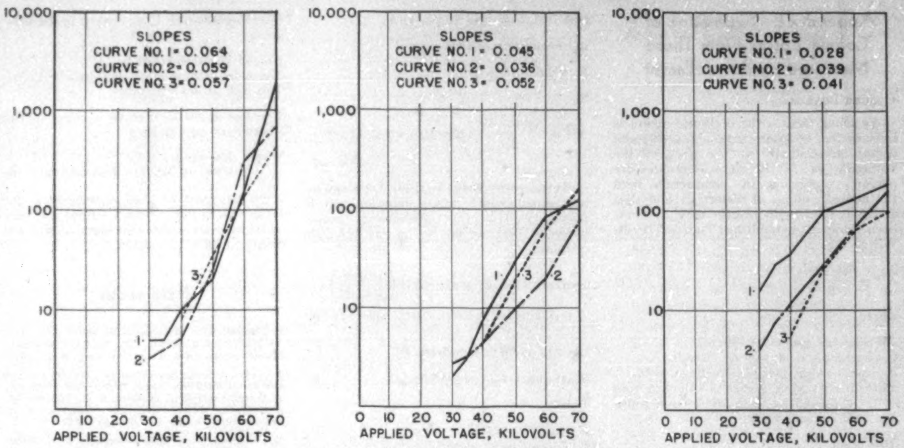
Although there are not sufficient data given to make complete calculations of ionization loss from the curves of Fig. 9, the average slope of the semilogarithmic loss curve can be calculated from the increase in pf with voltage above the break-point in each of the three curves, curves 1, 2, and 3, by methods similar to those applied to the curves of Fig. 7. These slopes can then be compared with the average slopes of the corresponding curves of Fig. 8. The comparison, given in Table III, shows that the slopes of the corona-intensity curves and the corresponding ionization-loss curves are substantially the same. In these three instances, as before, the two methods are equally effective in measuring corona discharge.

Since this analysis shows that the change in pf is a measure of the corona discharge, the discrepancy between the ionization loss as calculated from the corona-intensity measurements and that calculated from the pf measurements is unexplained. It will be noted that the ionization loss at 40 kv in curve 3 of Fig. 12 is less than the value calculated in the paper, equation 8, because of a difference in the methods of computing the increase in pf attributable to corona discharge. However, it is still much greater than the value computed from the corona-intensity measurements, equation 5. This suggests a need for more investigation of the quantitative interpretation of corona-intensity measurements.[1]

The choice of one of the three methods for either production testing or field checking can be based on considerations other than effectiveness in detecting corona discharge. From the standpoint of the user, engaged in acceptance and maintenance testing, a single test is to be preferred and the pf test is the logical choice because of the other undesirable conditions which it reveals.

REFERENCE

1. The Insulation of Electrical Equipment (book), **Willis Jackson,** editor. John Wiley and Sons, New York, N. Y., 1954, pp. 328–30.



Figs. 11, 12, and 13. Corona intensity, radio-influence voltage, and ionization loss versus voltage. Temperature: for Fig. 11 (left), 9 C; for Fig. 12 (center), 26 C; for Fig. 13 (right), 59 C

Curve no. 1: Corona intensity, coulombs $\times 10^{-12}$

Curve no. 2: Radio-influence voltage, microvolts

Curve no. 3: Ionization loss, milliwatts

G. M. L. Sommerman (Westinghouse Electric Corporation, East Pittsburgh, Pa.): These results of corona or ionization intensity in a specific type of insulated structure are very informative, especially since results with three different measuring methods are reported as a function of test voltage. The purpose of this discussion is to explore possible ways of reaching better quantitative agreement between the results of the three test methods, and this is not necessarily limited to bushings. A still better analysis could be made if the authors would provide data in terms of voltage stresses in the insulation (e.g., radial gradient at the center conductor) corresponding to one or more test voltages. Also, it would greatly add to the value of the paper to report ionization-intensity results obtained with all three methods before and during the accelerated life tests on bushings; see Table I.

The measurement of ionization losses in insulation by the pf versus voltage method has been used for many years on power cables. For cables, the increment in room-temperature pf on changing the voltage stress from about one-half to about double the operating value has been found to give in many cases a reliable indication of voids in cable insulation. This pf increment, designated as ionization factor, may be less than 0.0002 for new well-impregnated cable insulation. Tests of this type also show that the voltage increment of "solid" pf (pf not due to ionization) is less than 0.00005 at 125 volts, rms, per mil for dry, impregnated insulation, stress normal to laminations. For the particular laminated-core structure for which test results are given in Figs. 4 through 7, the authors conclude that solid pf increases with voltage even at voltages less than 20 kv (less than 30 volts per mil). Perhaps this is because the insulation contained moisture or other volatile constituents; these are recognized causes of positive values of $d\delta/dE$, and their effects are greater when the electric stress is not normal to laminations.

From Fig. 7, sloping pf voltage curves representing solid loss may be plotted by extrapolating the 0 to 20-kv portions of the measured curves. If these are subtracted from the measured curves, and plotted against the test voltage, we obtain curves of ionization factor versus voltage which are probably better indicators of ionization losses than are the total pf increments. Thus, for 26 C, the estimated ionization factor is found to be 0.00005 at 40 kv (compared with the value 0.00029 used by the authors in equation 28) and 0.00047 at 60 kv. Corresponding values of computed ionization losses are 4 megawatts (mw) at 40 kv and 75 mw at 60 kv. The figure of 4 mw is much less than the 20.5 mw computed from the total increment (see equation 28), and so is in much better agreement with the value 0.23 mw obtained by the authors in equation 5 from resonant-circuit corona-detector results. If the authors had chosen to compute their results at 60 kv for comparison of the two detection methods they would have obtained a better agreement because the corona intensity is over ten times greater at 60 kv. The computed ionization loss from the corona intensity would be at least 3.6 mw (assuming at least the same number of corona pulses per half-cycle) compared with the 75 mw just given for the corrected ionization-factor method.

The degree of agreement shown in this discussion, while better than indicated in Appendix I of the paper, is still far from satisfactory. We may examine the method of calculating the ionization loss from corona intensity. The basic equation, equation 1, is correct (I derived it independently before seeing this paper). One of the main things open to question is the use of only 8 pulses per half-cycle. If, in an insulated structure comprising impregnated-fibrous material of sizeable wall thickness, there are enough voids large enough to cause nearly 8 large pulses per half-cycle, it is most probable that there are a much greater number of smaller voids that will cause a large number of smaller pulses per half-cycle, especially as the voltage is increased moderately. While the task of integrating all of these pulses per half-cycle in a practical measurement may not be easy, it will have to be accomplished before really accurate ionization-loss values can be computed for functional insulated structures from corona-intensity measurements of the resonant-circuit type. Another point to be mentioned on corona-intensity measurements is that the ΔE required for equation 1 is as defined in the paper below the equation, and this value will be greater than the measured value obtained from the circuit shown in Fig. 3. This is because part of the test-piece voltage increment ΔE will not appear across the high-pass filter coupling inductor because part of ΔE will appear across the low-pass filter inductor in series with the test piece. Thus, quantitative comparisons with other methods cannot be made unless this condition is corrected.

Table III. Slope of Semilogarithmic Plots of Corona Intensity and Ionization Loss Versus Voltage Data from Figs. 8 and 9

	Slope	
Core No.	Corona-Intensity Curve	Ionization-Loss Curve
1	0.054	0.051
2	0.049	0.053
3	0.081	0.071

l, 74
1
55

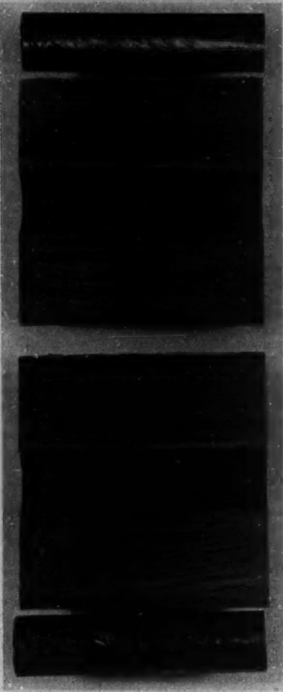


Fig. 14. Section of bushing removed because of television interference showing interfacial bombardment next to stem

In regard to the third method (radio-influence voltage), it appears that we are even farther from a quantitative ionization-loss measuring device. A small improvement could be made by multiplying the measured radio-influence voltage by the capacitance of the test piece (and any parallel capacitance) to obtain a radio-influence charge. This is analogous to the procedure followed in resonant-circuit corona-detection calculations. This would also make possible more accurate qualitative comparisons between results for test pieces having different capacitances, or results from laboratories using different parallel capacitances. If, as an example, we calculate values for 40 kv and 26 C, the radio-frequency influence charge is 5×10^{-4} micromicrocoulombs compared with 8 micromicrocoulombs, resonant-circuit value. Computation of the ionization loss could be made by the procedure using equations 1 through 5 if we knew the number of pulses per half-cycle, or could measure the integrated value thereof. It is obvious that this method measures only a small percentage of the voltage ΔE developed across the test piece. Other factors influencing the results for ionization-loss calculations from radio-influence results are coupling losses to the receiver, bandwidth frequencies, output meter response, etc.

W. F. Dunkle (Pennsylvania Power and Light Company, Hazleton, Pa.): This paper sets forth the problem associated with resin-bonded stem bushings. Ionization, if present at all, on paper insulation and continued sufficiently long can result in its failure. This phenomenon, when caused by voltage gradients on the paper surrounding voids in a bushing stem, results in growth of the void while in service because of interfacial bombardment. The only war to preclude this condition is to build a homogeneous solidly insulated stem completely devoid of any paper separation or voids.

For 20 years or longer the solid paper stem bushing has posed the problem of voltage gradient stresses. Over that period many bushings showing paper separation have been encountered. Some were repaired but many more were scrapped. Repairs were accomplished by forcing a very-high-grade synthetic varnish under pressure through the separation until it exuded from the other end and was then baked at 90 C.

Many 73-kv bushings of the type referred to in this paper have been cut up and examined to determine the cause for corona and high losses on stems. Fig. 14 shows one of the cut up bushings which was removed because of high noise level and it shows the cause for the interference and possibly ultimate failure. Interfacial bombardment is evident, showing imbedding of the minute copper particles in the paper itself. Fig. 15 shows variation of radio-influence level with bushing pf on a group of bushings which were removed because of interference.

Later models of the type-L bushing have been provided with an insulated ground sleeve which facilitates testing the cores' basic insulation. For all intents and purposes it is a 2-condenser layer-type bushing. The use of this feature can result in less capacitance being involved during the pf test than when the bushing is tested over-all to the ground flange. It facilitates interference measurements which eliminate effects of surface and cap to ground flange parameters. This results in higher pf for the bushing than is the case where the cap-to-ground flange measurement determines the pf. This may be erroneously construed to be due to ionization.

While on the subject of bushing ionization it is appropriate to touch on bushings of the bulk type. They are due for some criticism. They have been found with considerable quantities of water in the bushing particularly, when installed on circuit breakers. Venting of the bushing on a breaker during closure and condensing the entrapped moisture in the cool or light load period results in an accumulation of moisture in the lower end of the bushing. As much as a pint of water has been removed from bushings of this type. The remedial measure is to vent the bushing lower end. An example of such ionization is illustrated in Fig. 16.

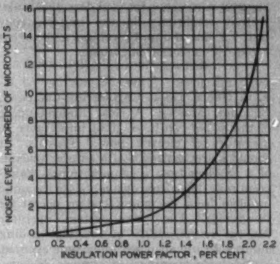


Fig. 15. Curve showing increase in interference with rise in bushing pf. 40-kv pf; 73-kv 400-ampere type-L bushings

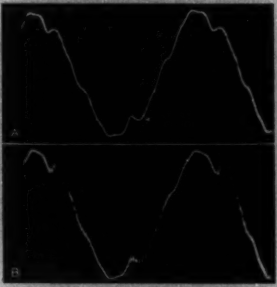


Fig. 16. Cathode-ray oscillogram showing ionization between stem and inside of 15-kv porcelain bushing at ground flange

A—Critical voltage 9,650 volts
B—Corona at 15,000 volts

Improper centering and rough internal bushing surfaces in the bulk bushing group result in ionization at less than operating voltage, the creation of copper compounds, very high pf, and frequently radio or television interference.

Electric utilities operating in fringe television and frequency-modulated radio areas are plagued by relatively low noise level apparatus creating interference. It is not only difficult but costly to locate the offending bushings and it is still more difficult to find one as a replacement which has no interference. Propagation of the interference can be for considerable distances.

The cause of the ionization, its location, the age of the bushing and its life history are items over which the manufacturer has little control. A bushing with an influence level of 100 microvolts in 1 year's time may or may not increase to 1,000, 5,000 microvolts or more. Evidence given in one of the other discussions is indicative that this can and does occur.

If the matter of paper separation cannot be overcome so that ionization at 110-percent rated voltage is zero, a redesign on another principle is necessary. Elimination of paper separation or voids is extremely difficult. High-grade paper saturated with oil is a suggestion to eliminate the voids.

Two fields remain to be explored in the manufacturer's laboratory: 1. Use of the electron microscope for void detection under ionized conditions on the low-frequency test. 2. Use of impulse voltage and current waves for detection of corona.

F. C. Doble (Doble Engineering Company, Belmont, Mass.): A clear understanding of this paper can be of great importance to operators of electric power systems. It should be recognized that a sincere analysis of a problem from a manufacturer's standpoint may differ from an analysis of the same problem from the power-system operator's standpoint. The power-system operator has the responsibility of keeping bushing failures to an absolute minimum even at considerable maintenance expense, because the over-all costs of a system outage resulting from a single bushing failure may greatly outweigh the mere replacement cost of a bushing. The author's presentation of test data can be of great assistance to power-system operators in evaluating the suitability of solid-core 69-kv bushings for service, provided certain points are clarified to make sure of a proper interpretation.

In the second paragraph the authors state "Most responsible writers on resin-bonded paper dielectrics recognize that the finished products may have air inclusions. If voids are located in areas of high electric stress, harmful levels of ionization may occur at operating voltages." Since resin-bonded paper dielectrics have this inherent weakness, the power-system operator is faced with the need for possible extra field testing equipment to cope with this special condition.

There are two ways of improving this design situation. As pointed out in the paper, one way is to reduce the possible voltage stress on any given void by the use of a multiplicity of conducting layers for stress control. Another approach would be to substitute for the resin-bonded paper an oil-impregnated paper as employed in bushings for higher voltages. It is our opinion that the maintenance testing economy indicates that one or both of these two alternatives should be considered at least at the 69-kv level. A further analysis of the paper appears to substantiate this opinion which is based on long experience with the maintenance testing of solid-core resin-bonded paper bushings.

In this paper extensive laboratory data are shown to prove the authors' conclusions that 1. the single-equalizer type of construction is "entirely satisfactory for bushings rated 69 kv and below," and "will last indefinitely in service if properly evaluated initially," and 2. "A method of corona detection when properly interpreted is a dependable quality control test." These statements connote that an initial low test reading on the factory production line will guarantee the future low reading and good behavior of a bushing under service conditions. Field test experience indicates that this optimism of the manufacturer cannot be fully shared by the operator. For example, 69-kv bushings of the type under discussion which had been in service were tested in the field and measured less than 100 microvolts radio-noise-influence voltage at operating voltage to ground (an accepted value in a factory test). When tested again, 1 year later, some of these same bushings had increased to more than 5,000 microvolts. This indicates that an initial low value does not guarantee future stability of this type of bushing.

Further, as a matter of record we believe that the authors have made some statements concerning pf tests which may be misinterpreted. They state that pf versus voltage measurements are not a reliable indicator of internal corona in this type of bushing. We believe this to be true as far as low-energy corona is concerned, but our tests have indicated that high-energy corona can be detected not only by a pf versus voltage test but in general even by a single pf test at 10,000 volts.

We would also like to point out that the large majority of bushing faults result from conditions other than void ionizations. Most of the faults are caused by moisture entrance, carbon deposits, mechanical faults, etc. Experience has demonstrated the ability of the pf test at voltages up to 10 kv to detect such faults before failure even in bushings built for 220-kv operation.

Extensive field testing experience records appear to indicate that the operating voltage classification for a solid-core main insulation (single-equalizer type), particularly of the resin-bonded paper construction, should be reduced below the 69-kv level.

H. H. Brustle, D. L. Johnston, and **A. L. Scheideler:** We appreciate the detailed interest which Mr. Wickham and Mr. Sommerman have shown in the data presented. It is believed, however, that several comments are warranted by their analysis of the pf data. No justification can be found for the assumption that bushings exhibiting linear increases in pf with voltage are corona-free, while exponential increases are due to ionic discharges in the core material. If this were always the case then Mr. Wickham's analysis as applied to curve 1 of Figs. 8 and 9 would show that ionic discharges start at 25 kv whereas the resonant-circuit method, which is a much more sensitive corona detector, shows ionic discharges starting at 50 kv. Another difficulty in such an approach is shown by the fact that Mr. Sommerman assumed 0 to 20 kv as the linear region while Mr. Wickham chose 0 to 35 kv. Actually, the rate of increase in pf with voltage caused by ionic discharges depends on the distribution of void sizes and void locations.

The usefulness of pf tests in revealing the presence of high-intensity corona at levels above operating is attested to by both Mr. Wickham and Mr. Sommerman and corroborated by our corona and radio-noise results. However, in the very important region near operating voltage, direct measurements of discharge phenomena such as corona and radio noise are much less subject to misinterpretation. There is some question, e.g., as to whether Mr. Wickham would have drawn the same conclusions with reference to the 59 C data of Fig. 7 in the absence of corroborating radio-noise and corona data. All points at 59 C fall on a straight line to within 0.01-per-cent pf. This is beyond the precision of most field testing pf equipment. The data given in the paper were obtained on a bridge with a precision of greater than 0.001 pf.

It is shown (ref. 1 of Mr. Wickham's discussion) that pf can actually decrease with voltage. This is generally attributed to carbonized voids or tracking. We have, however, observed similar phenomena on developmental cores which have never been subjected to voltage stress.

We fully appreciate the value of pf as an evaluation tool in field testing. Our feeling is, however, that its greatest utility in the field lies as a means of measuring changes with time in service. It is, of course, of great value in the factory as an indicator of pf level and of changes during factory-proof testing.

We well recognize the need for a true corona pulse integrating device as Mr. Sommerman mentions. Since no such device is available at the present time, our estimates of pulse rate were based on experiments using the pulse counting technique. Mr. Sommerman's point about many smaller pulses being present is well taken and should be considered. A calculation made considering these smaller pulses indicates that approximately 10,000 pulses per second of 1-micromicrocoulomb intensity are necessary to affect pf measurement as much as 1,000 pulses of 8 micromicrocoulombs as used in the paper. Even then the total ionization loss is a small percentage of the dielectric loss calculated from pf measurements. At the higher voltages, the pulse count of 1,000 pulses per second as used in the paper would, of course, not hold because the increased stresses throughout the core structure would have caused many more voids to ionize. Therefore, better agreement should be expected at the higher voltages where corona is significantly contributing to the pf measurement.

The method used for measuring ΔE was that described in reference 2 of the paper. This method takes into account the fact that the theoretical ΔE as measured across the test piece and the ΔE as measured across the coupling inductor are very much different. Because such a calibration was made, and the theoretical ΔE used, calculations made in the paper are justified.

Mr. Dunkle's data are a valuable contribution to complete understanding of what caused the deterioration in a small percentage of the single-equalizer design at 69 kv. He verifies our statement that field deterioration is radial and not axial in nature. Work to improve stress conditions at the center conductor is in process and preliminary results are very encouraging.

Two ways of limiting voltage stress concentrations are mentioned by Mr. Doble. A third way is exemplified by the simple design shown in Fig. 1 where as few stress equalizers as practicable are employed. Since the electric stress tends to concentrate at electrode terminations, the design with fewer electrodes has less areas of stress concentration.

Mr. Doble states that the failure of even a single bushing is serious to the operator. We must insist that such a failure is considered seriously by the manufacturer as well. The advice and experience of operators such as Mr. Dunkle, Mr. Wickham, and Mr. Doble will be needed to assess fully the value of radio-noise measurements as a tool for field evaluations.

Improving Field Analogues Through Conformal Mapping

H. K. FARR
MEMBER AIEE

W. A. KEEN, JR.
ASSOCIATE MEMBER AIEE

Synopsis: Conformal mapping leads to improved models for the electrolytic tank and similar analogues of fields encountered in engineering problems. The effect of tank boundaries can be eliminated and regions of special interest can be greatly magnified. Conformal mapping has a validity more general than is commonly realized. It can, for example, be applied to the construction of models of rotationally symmetric fields.

A CONDUCTION analogue such as the electrolytic tank is a familiar device for the solution of problems involving electric, magnetic, thermal, and other fields. The accuracy of measurements made in this way is often limited by distortion of the field produced by the boundaries of the conducting medium and by practical limitations on the size of the model. In the design of a high-voltage bushing, e.g., it may be desired to study the electrostatic field by means of a model in the electrolytic tank. If the model of the bushing takes up a large proportion of the tank area, the walls of the tank will cause a distribution of voltage which may differ considerably from that around an actual bushing in free space. On the other hand, if the model is made very small in comparison with the tank, model construction and plotting become more difficult and meniscus effects in the electrolyte distort the field.

Boothroyd, Cherry, and Makar[1] described an ingenious double tank which was used for the study of impedance functions and which essentially eliminated the usual errors caused by the tank walls. In its original form, the double tank could be adapted to field problems but only to those of translational symmetry, i.e., 2-dimensional fields of the type that can be represented in a shallow tank with a uniform electrolyte of constant depth, e.g., the electrostatic field of long straight parallel conductors.

Actually, the technique used by Boothroyd et al. is a special case of a more general method based on conformal mapping. This more general method permits the elimination of boundary errors appearing in conduction analogues of electrostatic fields in either uniform or composite dielectrics and of either the simple 2-dimensional type or of the type represented by figures of revolution, i.e., fields of so-called translational and rotational symmetry respectively. Furthermore, the tank need not be circular in outline. In some problems the desired results may be obtained in a single tank. Also, for either translational or rotational symmetry, the technique may be applied to either direct or conjugate analogues. It is also possible to enlarge greatly a small region of particular interest while keeping the rest of the model small and therefore realize the highest accuracy where it is needed.

These techniques apply to thermal, magnetic, and other fields as well as to electrostatic fields. However, for simplicity only the latter will be discussed. The extension to fields of other types follows on replacing potential and dielectric constant by analogous quantities such as temperature and thermal conductivity or flux function and reciprocal of permeability. These analogies are discussed in an earlier paper.[2]

Conformal Mapping

The techniques to be described depend on an extension of the use of the technique of conformal mapping. This is a method which has been widely used for the computation of fields but which has not received much attention in connection with analogue methods of field study.

A potential field in two dimensions may be described by expressing the potential in terms of the position in rectangular coordinates x and y

$$\phi=\phi(x,y)$$

It is often useful to transform the shape of one field into another configuration which is more easily determined. Each point (x,y) of the original field is imagined as corresponding to another point (u,v) at the same potential in the new field. That is

$$\phi(x,y)=\phi'(u,v)$$

Such a transformation is useful only if the new field is an actual physically realizable electrostatic field. It is well known that a transformation known as conformal will do just that for the simplest 2-dimensional fields—those satisfying Laplace's equation. An electrostatic field in a uniform dielectric is such a field.

In connection with conduction analogues, it is useful to extend conformal mapping to fields of a more general type, namely, to electrostatic fields in dielectrics where the permittivity may vary with position and to analogues of such fields. In doing so, the only requirement to be observed is that the permittivity (or its analogue) remain invariant under the transformation. The validity of this generalization is proved in the Appendix. It is possible that this theorem has already appeared in the technical literature, however, the authors have not come across any reference to it.

The analogue of a translationally symmetric electrostatic field with varying permittivity is a conducting sheet in which the conductance per square varies in proportion to the permittivity of the initial electrostatic field. Practically, this may be realized in a very shallow tank where the depth of electrolyte is varied in the required manner. It is therefore possible to construct a model representing a conformal transformation of a field provided that the depth of electrolyte in the transformed model is the same at any point as it would be at the corresponding point of a model which represented the field directly. When the depth of electrolyte is varied smoothly, a very shallow tank must be used to approximate a thin sheet. Where there is an abrupt step, a row of conducting pins will insure a good approximation to the correct field.

So far, the authors have found it convenient to use one particular conformal transformation exclusively. This is an inversion with respect to a circle. If ρ and ϕ are polar co-ordinates in the initial field and r and θ are those in the transformed field, the transformation is represented by

$$r=a^2/\rho,\ \theta=\phi$$

where a is a constant.

Fields associated with figures of revolution (rotational symmetry) may be represented by a conducting sheet whose conductivity varies in a prescribed manner

---

Paper 55-105, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 19, 1954; made available for printing December 9, 1954.

H. K. Farr and W. A. Keen, Jr., are with the General Electric Company, Pittsfield, Mass.

The authors wish to acknowledge helpful suggestions by E. S. Weibel for the section in the Appendix entitled "Differential Equation Method."

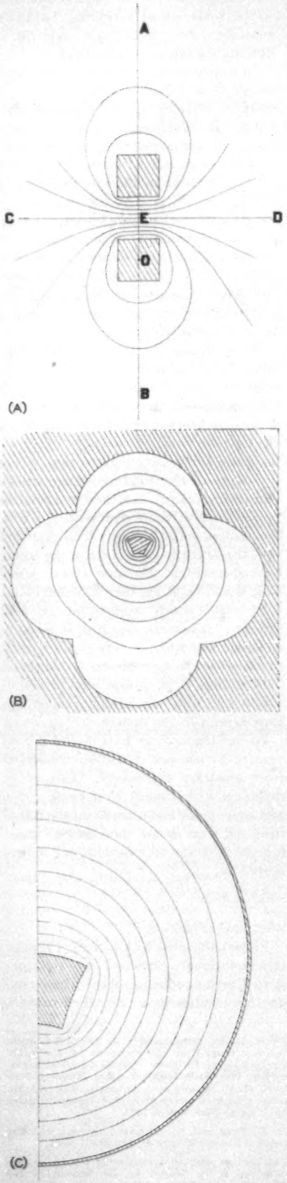


Fig. 1. Application of conformal mapping to the construction of a model for the electrostatic field between two conductors of square cross section

A—The original field spreads out to infinity in all directions. However, inversion in a circle centered at O leads to

B—Where the entire field is enclosed by the outer electrode

C—Exploiting symmetry permits further simplification

as a function of radius. Such fields can, therefore, be regarded as a special case of translationally symmetric fields of varying permittivity; hence, they are also susceptible to a conformal transformation, in particular, inversion in a circle.

In many cases, it is a simple matter to translate the information obtained from the transformed model back to the conditions of the initial field. Although a complete field plot must be transformed point by point, this is often not required. Thus the conductance between a pair of electrodes (analogous to capacitance) remains unchanged by a conformal transformation and gradient values need only be multiplied by the factor r^2/a^2 to transform them back to the initial field.

Elimination of Tank Boundary Errors

An important aspect of conformal mapping for conduction analogues is its ability to transform a field of infinite extent into one of finite extent which can then be represented in a tank or on a conducting sheet of practical size. By means of an inversion in a circle, the entire region outside a circle of radius a is mapped inside the same circle while the region originally inside goes over to the outside. Hence, if the circle of inversion is chosen to lie inside one electrode, the original field will lie entirely outside this circle and, therefore, be transformed into a finite area. Actually, it is not even necessary for the circle to lie inside the electrode as long as its center does.

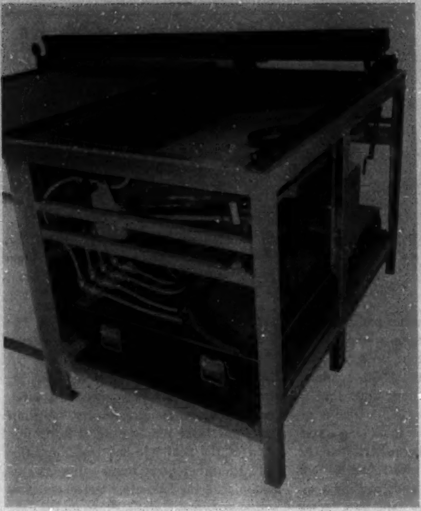
A simple application of this method is illustrated in Fig. 1. This concerns the electrostatic field caused by two parallel conductors of rectangular cross section (shown hatched). The contour lines in these figures are equipotentials drawn at intervals of 10 per cent of the total voltage.

An attempt to plot the field of Fig. 1(A) directly in an electrolytic tank would lead to errors from the tank walls. However, inversion in a circle produces the configuration of Fig. 1(B) which avoids this difficulty. The lower electrode of Fig. 1(A) is transformed in Fig. 1(B) into the outer hatched area. This area is bounded at the inner edge by four semicircles which form a roseate contour that completely encloses the electric field. An electrode of this shape forms the outer boundary of the model. The upper electrode of Fig. 1(A) is transformed into the small inner hatched area of Fig. 1(B). Practically, it is usually quite easy to lay out the electrode shape required for a transformed model. Inversion in a circle belongs to the so-called bilinear type of transformation which transforms circles into circles, provided a straight line is included as a special case of a circle.

For a symmetrical configuration like that of Fig. 1(A), it is possible to simplify the problem further by exploiting the symmetry axes. Thus, it is known without plotting the field that the straight lines AB and CD are respectively a flux line and an equipotential line. Hence, the model need represent only one quadrant, say AED, of the original field. If the center of inversion is chosen anywhere outside this quadrant, the lines AB and CD will be transformed into a closed contour which will constitute the boundaries of the model. Such a plot is shown in Fig. 1(C) where the outer semicircular electrode represents the equipotential line ED and the insulating tank wall along the diameter is the transformation of the flux line AE. Although the same center of inversion was used to obtain both Figs. 1(B) and (C), it would have been possible in the second case to choose a point outside the electrode if that were desirable. Actually, Fig. 1(C) is identical with the inner right-hand portion of Fig. 1(B). The 50-per-cent contour line in Fig. 1(B) is precisely a circle whose right half is the outer boundary of the model shown in Fig. 1(C).

Preferential Scaling

The proper choice of transformation will lead to a model which represents the region of interest on a favorable scale and suppresses other regions. This is done by taking the center of the circle of inversion at a point as close as possible to the interesting region but still within the electrode. This can be very useful for the detailed examination of a small area even though no transformation is needed for boundary effect errors. If, e.g., the field close to one of the corners of the lower electrode of Fig. 1(A) were the point of interest, the center of inversion O could be moved up into that corner, still remaining within the hatched area. Then Fig. 1(B) would take on a somewhat different appearance; most of the area



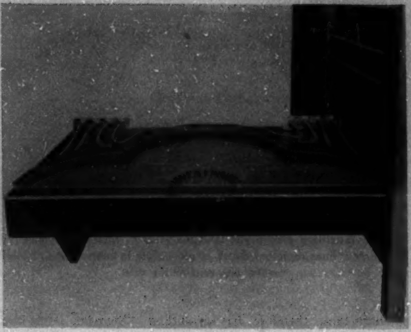


Fig. 2 (left). Auxiliary (lower) tank in position connected to primary sloping tank

Fig. 3 (above). Auxiliary tank. Cables connect outer row of pins to primary tank. Inner semicircular resistance grid supplants deepest portion of electrolyte

would be occupied by portions of the original field near the corner of interest and all the remainder of the original field would be squeezed into a small area at the center.

The Double-Layer Model

A conformal transformation will often transform one field into another which can be represented with much greater accuracy than the initial field. In some problems, however, it may be desirable to retain the region of interest in its original form. Also, the requirement of taking the circle of convergence inside one electrode may lead to unfavorable scaling; in the case of point electrodes, as used in the study of impedance functions, it is, in fact, out of the question.

By the use of a double-layer model it is possible to realize the advantages of the conformal transformation without these defects. The double tank used by Boothroyd et al.[1] is the simplest example of this technique. The initial field is divided into two portions by the circle of inversion. The portion inside the circle is left unchanged. The portion outside the circle is subjected to the usual transformation which results in a region that, like the first portion, just fills the area within the circle of inversion. In this way the entire area of the initial problem, including all regions inside or outside electrodes, can be represented by two circular regions. In the double tank these become two disk-shaped layers of electrolyte placed one above the other and connected along their edges. The upper layer is ordinarily the untransformed portion of the field. The two layers of electrolyte are contained in a single circular tank and separated by a horizontal insulating disk. The diameter of the disk is slightly smaller than that of the tank to permit conduction around the outer edge; or an array of electrodes is placed around the boundary.[3] The periphery of the tank is the circle of inversion. The connection along the edges of the disks provides the correct boundary conditions between the two regions.

This technique has been adapted to a very simple model in which the electrolyte is replaced by conducting sheets. The sheets consist of Teledeltos facsimile paper, which is similar to that used in a simple field plotting device described by Helm and Hand.[4] Two conducting disks are separated by a thin cardboard disk of slightly smaller diameter. The two outer disks are connected together along their edges by a conducting adhesive. The conductivity of the adhesive must be high enough to prevent any appreciable voltage drop from one disk to the other but low enough to avoid any shunting effect in the peripheral direction. Although not susceptible to as high an accuracy as electrolytic models, the double-layer conducting paper model has the advantage of convenience and simplicity.

Rotational Symmetry

The authors have constructed a double tank to represent fields of rotational symmetry. As already pointed out, fields of this type may be subjected to a conformal transformation. It is therefore possible to adapt to this situation the double-layer tank for fields of translational symmetry. For rotational symmetry, however, the depth of electrolyte is not constant and this makes it necessary to use two separate tanks. An array of small plate electrodes was placed around the boundary of each tank and each electrode in one tank was connected to a corresponding one in the other tank. The so-called primary tank is just the conventional sloping tank commonly used for the direct representation of fields having rotational symmetry. This comprises the portion of the initial field which is plotted without transformation. The remaining, transformed portion appears in the auxiliary tank where the depth of electrolyte varies in a somewhat more complicated fashion, as determined by the rule that the conductance per square must be invariant under the transformation. With this arrangement, it is unnecessary to use circular tanks. A rectangular shape has advantage where space is limited or where one of the insulating boundaries of the model is to be a straight line. This shape was therefore chosen for the primary tank where the more interesting region of the field can be represented without distortion. The

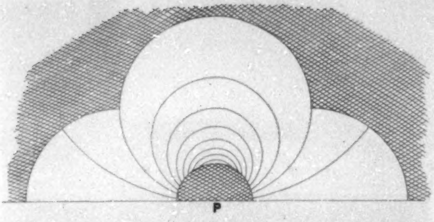


Fig. 4. Auxiliary electrolytic tank for fields of rotational symmetry. The contour lines represent depth of electrolyte in intervals of 1/2 inch starting with zero on the axis

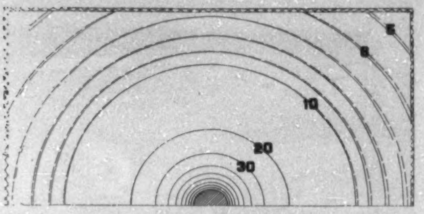


Fig. 5. Electrostatic field between concentric spheres. Solid lines are equipotential plot in primary tank, broken lines computed. Outer sphere was placed in auxiliary tank

auxiliary tank, therefore, has an outline which is obtained by transforming the rectangular outline of the primary tank. For simplicity, the center of inversion was taken on the axis of revolution. Fig. 2 shows the two tanks in position one above the other. The upper one is the primary tank placed so as to be accessible for field plotting. The auxiliary tank is shown more clearly in Figs. 3 and 4. As the axis of revolution remains a straight line after transformation, the shore line of the electrolyte in the auxiliary tank is a straight line. The depth of the electrolyte increases continuously as one proceeds from the axis into the tank along a semicircular path passing through the point *P*. The depth of the electrolyte in the primary tank may be represented by contours, i.e., level lines parallel to the axis and uniformly spaced. To construct a contour map showing the depth of electrolyte in the auxiliary tank, it is merely necessary to transform the contour lines from the original plane in the same way that a field pattern would be transformed. Since the contours in the primary tank are straight lines parallel to the axis, the resulting circles are all tangent to the axis at the point *P* representing infinity.

Close to the point *P* the depth of the electrolyte would approach infinity. Aside from the practical difficulty of making the electrolyte very deep, an error is introduced where the depth changes very rapidly. Ideally, the conducting medium would be a thin sheet of varying conductance per square. This can be approximated successfully by an electrolyte of uniform conductivity but variable depth as long as current flow is essentially horizontal. But the approximation is poor when the bottom slopes sharply. For this reason, a network of lumped resistances was used to replace the electrolyte in this area.

The conductance per square of the electrolyte next to the boundary of one tank must be the same as that at the corresponding point in the other tank. To reduce the size of the region where the resistance network must be used, the conductivity of electrolyte in the auxiliary tank was chosen to be four times that in the primary tank. The depth then is reduced by a factor of four in passing across the boundary from the primary to the auxiliary tank. The electrolyte in the auxiliary tank then extends the linear dimensions of the primary tank by a factor of about five and the resistance network can be made fine enough to give a good representation of the remaining region out to infinity.

Results

Checks made by using the auxiliary tank or a single tank with transformation have given good agreement with theory on those configurations tested which were susceptible to calculation. Fig. 5 shows a plot of the portion of the field in the primary tank corresponding to two concentric spheres whose radii are in the ratio of 60 to 1. It is evident that, had it been necessary to place both electrodes in a single tank, the inner one would have been so small as to result in gross errors. Since the tank has a sloping bottom, the depth of electrolyte is very small around the tiny inner electrode. Meniscus effects and irregularities of the tank bottom are, therefore, most serious. With the help of the auxiliary tank, however, the field in the primary tank reproduces the theoretical one with good fidelity from the inner sphere right up to the edge of the tank. The error in potential along the tank boundary nowhere exceeds about 0.2 per cent.

Conclusions

A simple conformal transformation such as an inversion in a circle will provide greatly improved accuracy for conduction analogues of many two dimensional fields. This technique may be applied in two ways: a portion of the field which extends to infinity may be reduced to finite size by a single transformation; or the entire field may be mapped onto two interconnected conducting layers of finite size, one layer representing a finite portion of the field without transformation and the other layer representing the remainder of the field after transformation.

Conformal transformations may be applied to models of simple 2-dimensional fields and to fields having symmetry of rotation. It may be used even though the material medium is not uniform (variable conductivity, permittivity or permeability). It may be applied to both direct and conjugate analogues of such fields. The conductivity at any point must have the same value before and after transformation.

These principles have been used successfully in a double electrolytic tank for fields of rotational symmetry to solve problems in electrostatics which were previously intractable. A double layer model for plane 2-dimensional fields may be constructed very simply from two circular sheets of conducting paper bonded at the edges with a conducting cement.

Appendix. Validity of Conformal Transformation

In this section the validity of the conformal transformation for regions of varying permittivity is demonstrated and the double tank is justified. In a qualitative way, it might be expected that the first follows from the fact that the capacitance of a configuration is invariant under a conformal transformation which conserves permittivity. Such a transformation should not disturb the ratio of flux to voltage in any small segment of the field. For a more rigorous demonstration, two alternate proofs are supplied. The first is based on a consideration of a field integral, whereas the second analyzes the behavior of a differential operation common to several important fields.

Variational Method

A particularly elegant demonstration of the generalized conformal transformation was suggested to the authors by Dr. H. Poritsky of the General Electric Company. It is based on the theorem that the field energy is a minimum, i.e., the potential $\phi(x,y)$ is a solution of the equation

$$\delta\int_s \epsilon\left[\left(\frac{\partial\phi}{\partial x}\right)^2+\left(\frac{\partial\phi}{\partial y}\right)^2\right]dx\,dy=0$$

In a conformal transformation which leaves the permittivity ϵ invariant, the field energy for any element of area remains invariant.[5] Hence, the value of the foregoing integral will be invariant and, therefore, also the potential function which minimizes it. That is to say, a conformal transformation which leaves potential and permittivity invariant will result in a potential function which is the correct physical potential in the transformed co-ordinates. Of course, for given electrode potentials, the permittivity may be multiplied by a constant factor without changing the potential distribution and, similarly, for the conductance per square of an electrolyte. In the case of the auxiliary tank, however, it will be found that the conductance is fixed by the primary tank and by the boundary conditions arising from the connection between the two tanks.

Differential Equation Method

Here the expression of interest is

$$\omega=\nabla\cdot\epsilon\nabla\phi$$

This occurs in the equation

$$\nabla\cdot\epsilon\nabla\phi=-\rho \tag{1}$$

for the electrostatic potential caused by a charge density ρ. A similar expression occurs in the equation

$$\phi+\frac{1}{\alpha}\nabla\cdot\frac{1}{\mu}\nabla\cdot\phi=\phi_0 \tag{2}$$

where α, μ, and ϕ_0 are given functions of position which describes the 2-dimensional magnetic field of alternating currents in a conductor.[2]

In terms of the rectangular co-ordinates x_1, x_2, and x_3

$$\omega=\frac{\partial}{\partial x_i}\left(\epsilon\frac{\partial\phi}{\partial x_i}\right)=\frac{\partial\epsilon}{\partial x_i}\frac{\partial\phi}{\partial x_i}+\epsilon\frac{\partial^2\phi}{\partial x_i^2}$$

where summation is taken over the index i. In this and the following equations, the familiar convention of summing over any index which appears twice in the same term is used. Consider a transformation which takes the point x_1, x_2, x_3 into the point y_1, y_2, y_3, having the same values of ϕ and ϵ. If the transformation is such that ω has the same form in the new co-ordinates, then the differential equation is invariant under the transformation and ϕ will be the correct physical potential in the new configuration.

To make the transformation, one may use the coefficients

$$g_{ji}=\frac{\partial y_j}{\partial x_i}$$

Then

$$\frac{\partial g_{ji}}{\partial x_i}=\nabla^2 y_j$$

where

$$\nabla^2=\frac{\partial^2}{\partial x_i^2}$$

$$\frac{\partial\phi}{\partial x_i}=\frac{\partial\phi}{\partial y_j}g_{ji}$$

$$\frac{\partial\epsilon}{\partial x_i}=\frac{\partial\epsilon}{\partial y_k}g_{ki}$$

Also

$$\frac{\partial^2\phi}{\partial x_i^2}=\frac{\partial}{\partial x_i}\left(\frac{\partial\phi}{\partial y_j}g_{ji}\right)=\frac{\partial^2\phi}{\partial x_i\partial y_j}g_{ji}+\frac{\partial\phi}{\partial y_j}\nabla^2y_j=$$
$$\frac{\partial^2\phi}{\partial y_j\partial y_k}g_{ji}g_{ki}+\frac{\partial\phi}{\partial y_j}\nabla^2y_j$$

$$\omega=\left(\frac{\partial\phi}{\partial y_j}\frac{\partial\epsilon}{\partial y_k}+\epsilon\frac{\partial^2\phi}{\partial y_j\partial y_k}\right)g_{ji}g_{ki}+\epsilon\frac{\partial\phi}{\partial y_j}\nabla^2y_j$$

For a conformal transformation in two dimensions, the Cauchy-Riemann equations give

$$g_{11}=g_{22},\ g_{12}=-g_{21}$$

This means that $\nabla^2y_j=0$ and that

$$g_{ji}g_{ki}=0 \text{ if } j\neq k$$

It further follows that, if $j=k$, the value of $g_{ji}g_{ki}$ is independent of j and, hence, may be designated by the indexless quantity g^2. It is then found that g is just the scale factor at any point

$$g=\frac{ds'}{ds}$$

where ds and ds' are elements of length before and after transformation. The fact that ds'/ds is independent of the direction of ds is characteristic of a conformal mapping. The expression for ω is now

$$\omega=\nabla\cdot\epsilon\nabla\phi=g^2\left(\frac{\partial\phi}{\partial y}\frac{\partial\epsilon}{{}_j\partial y_j}+\epsilon\frac{\partial^2\phi}{\partial y_j^2}\right)=$$
$$g^2\nabla'\cdot\epsilon\nabla'\phi=g^2\omega'$$

where the operator ∇ refers to the original co-ordinates and ∇' to the new co-ordinates, i.e., a conformal transformation merely changes the expression for ω by a factor of $1/g^2$. This leaves invariant the equation

$$\nabla\cdot\epsilon\nabla\phi=0$$

for the electrostatic potential in a charge-free region. Where charge is present, equation 1 becomes

$$\nabla'.\epsilon\nabla'\phi=-\rho/g^2$$

The charge density must be changed by a factor of $1/g^2$ to produce the same potential. Since the area changes by the factor g^2, the total charge in a given region remains invariant. Equation 2 for an a-c magnetic field becomes

$$\phi+\frac{g^2}{\alpha}\nabla'\cdot\frac{1}{\mu}\nabla'\phi=\phi_0$$

The parameter α must be changed by a factor of $1/g^2$ for the equation to remain invariant. In capacitance-conduction analogues for such fields, α is proportional to the capacitance per unit area. Hence, the capacitance for a given region is invariant under the transformation.

Gradient Transformation

It has been shown that the quantity $\nabla\epsilon\nabla\phi$ is multiplied by $1/g^2$ as a result of a conformal transformation. Likewise, the magnitude E of the gradient $\nabla\phi$ changes by $1/g$. That is

$$E=gE'$$

For inversion in a circle $g=r/\rho=r^2/a^2$, where ρ and r are the distances from the origin before and after transformation and a is the radius of the circle of inversion.

Boundary Conditions

The following conditions now hold for the conformal transformation: potential and permittivity are invariant; and the gradient changes by a factor g. As was pointed out previously, the invariance of ϵ is sufficient but not necessary. It suffices that $\epsilon'=k\epsilon$, where k is a constant. This more general condition merely means that ω transforms according to the equation

$$\omega=\frac{g^2}{k}\omega'$$

so that g^2 is replaced by g^2/k.

In the case of two electrolytic tanks (primary and auxiliary), the boundaries must be interconnected in such a way as to produce the correct conditions, namely

$$\phi=\phi',\ E_n=gE_n'$$

where E_n is the normal component of the gradient. In analogy to the condition on ϵ, it is also required that $\sigma'=k\sigma$ throughout both tanks, where σ is the conductance per square.

Interconnecting corresponding points of the two boundaries actually results in

$$\phi=\phi' \text{ and } dI=dI'$$

where $dI=i_nds$ is the current flowing across a small portion of the boundary of length ds, and i_n is the normal component of the current density. Hence, interconnection satisfies the correct condition on the potential and for the gradient gives

$$i_n=i_n'\frac{ds'}{ds}=gi_n'$$

or

$$\sigma E_n=g\sigma'E_n'$$

Hence, the required conditions are $\sigma=\sigma'$ along the boundary and $\sigma=k\sigma'$ everywhere. Both conditions are satisfied by choosing $k=1$ so that $\sigma=\sigma'$ everywhere.

References

1. AN ELECTROLYTIC TANK FOR MEASUREMENT OF STEADY STATE RESPONSE, TRANSIENT RESPONSE AND ALLIED PROPERTIES OF NETWORKS, A. R. Boothroyd, E. C. Cherry, R. Makar. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 96, pt. I, no. 99, May 1949, pp. 163–77.

2. SOME ENGINEERING APPLICATIONS OF THE ELECTROLYTIC FIELD ANALYZER, H. K. Farr, W. R. Wilson. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1301–09.

3. DESIGN OF ELECTRIC WAVE FILTERS WITH THE AID OF THE ELECTROLYTIC TANK, A. R. Boothroyd. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 98, pt. IV, 1951, pp. 65–93.

4. FIELD PLOTTING METHOD USES MODELS, S. Helm, A. R. Hand. *Industrial Laboratories*, Chicago, Illinois, vol. 4, no. 9, Sept. 1953, pp. 122–24.

5. ELECTROMAGNETIC FIELDS: THEORY AND APPLICATION. VOLUME 1. MAPPING OF FIELDS (book), Ernest Weber. John Wiley and Sons, New York, N. Y., 1950, 1st ed., p. 304.

◆

Discussion

John F. H. Douglas (Marquette University, Milwaukee, Wis.): This paper gives a solution for the difficult experimental problem of experimental mapping by means of a water trough when the field spreads indefinitely to an infinite distance. For the sake of completeness an alternative solution should be mentioned.

When a spreading field has, at remote points, the characteristics of a line doublet, the electrolytic trough may be built with a boundary conforming to that doublet. This is a boundary consisting of four circles alternately conducting and insulating and intersecting each other at right angles. A supplementary tank must then be used to imitate the geometric resistance of the field outside of the boundary of the tank.

I have used this principle to construct models from conducting paper of an Edison bipolar dynamo, the field due to a polarized cylinder, the field due to two parallel bus bars, and the field due to a thundercloud. One advantage of this technique, when it can be applied, lies in the fact that the field does not have to be remapped.

With regard to the row of pegs in a trough having a discontinuous depth, I should like to refer to chapter 2, pages 6 and 7, of my book "Electric Magnetic and Thermal Fields" (privately published).

H. K. Farr and W. A. Keen, Jr.: The technique described by Mr. Douglas offers another approach to the problem of fields which extend to infinity. The method is quite simple in that his supplementary tank has only two electrodes. In many practical problems it would certainly produce much more accurate results than a single conventional tank. It does, however, involve an approximation which is not required in the techniques based on conformal mapping and which would be more or less serious depending on the particular configuration. There is also some question as to whether the method could be extended to fields of rotational symmetry or composite dielectrics. Furthermore, it should be emphasized that when using an auxiliary tank of the type based on conformal mapping it is not necessary to remap the field in the primary tank. Remapping would be resorted to only where conformal mapping is done with a single tank or where it is desirable to know the field in the auxiliary tank as well.

Since submitting this paper we have received a private communication from Lars V. Ahlfors of Harvard University pointing out that the invariance of the equation $\Delta \cdot \epsilon \Delta \phi = 0$ under a conformal mapping is used in the field of pure mathematics. This is the property, which makes possible the more general techniques described in this paper, e.g., the conformal mapping of fields in composite dielectrics.

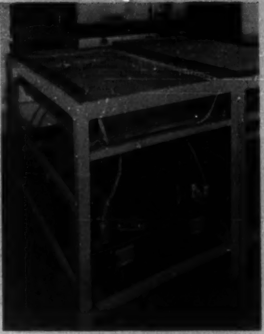


Fig. 6. Double tank for fields of translational symmetry, primary flat tank above

Since this paper was submitted, we have completed and successfully operated an auxiliary tank for use with a rectangular tank designed for fields of translational symmetry. This is shown in Fig. 6 which corresponds to Fig. 2 illustrating the tanks for fields of rotational symmetry. The center of inversion is in the center of each tank. For obvious practical reasons the auxiliary tank has been rotated 90 degrees with respect to the primary tank, i.e., the long axis of one tank actually corresponds to the short axis of the other.

Vol. 7
Pt. 1
1955

A Portable Telegraph Bias and Distortion Measuring Instrument

W. D. CANNON
FELLOW AIEE

IN THE modern telegraph system, signals are transmitted over various trunk facilities derived from wire lines or radio relay systems. The trunk telegraph channels furnish facilities between a number of switching centers located at strategic points throughout the country. These centers then distribute traffic to the local points of destination. In addition, various other services must be accommodated by suitable network configurations. These services comprise such arrangements as are involved in leased network circuits, private switching systems, etc. It is necessary that flexibility be maintained in this complex network of circuits so that circuit patching and routing may be accomplished without making undue alignment changes in the circuits themselves.

In the operation of a telegraph network, it is often necessary to interconnect telegraph circuits indiscriminately to form through-trunks and local distribution circuits. It is not only essential that circuits and equipment be compatible with each other but also that signals as measured from point to point do not exceed a permissable tolerance with regard to signal distortion. The signal quality as measured at interconnecting points permits prediction of the behavior of the interconnected circuit. With the aid of a signal distortion indicating device, adjustments can be made to provide the best possible circuit operation.

ol. 74
t. 1
955

Bias and Distortion

Bias is usually the most important of the several types of distortion which may be present in telegraph circuits. Bias results in a uniform displacement of like signal transitions so that all marking signal intervals are uniformly lengthened or shortened and all spacing signal intervals are correspondingly shortened or lengthened. In start-stop teleprinter operation, bias results in a uniform displacement of space-to-mark transitions with respect to the start pulse mark-to-space transition. For this reason, bias is sometimes measured by observing the time of occurrence of space-to-mark transitions with respect to the beginning of the start pulse. Since bias displaces the mark-to-space transitions uniformly, it is possible to measure bias not only with respect to the beginning of the start pulse but also with respect to the position of any other mark-to-space transition occurring in the signals. If the measurement is made of the space-to-mark transitions of single unit spacing signals with respect to their corresponding mark-to-space transitions, the bias indication is based upon the average duration of a spacing unit interval. The bias measurement is accomplished by measuring the average time interval occupied by single unit spacing signals in comparison with the time interval of a perfect unit length spacing signal. Measurements are made of spacing single unit signal intervals rather than of mark-signal intervals because in start-stop operation the duration of the marking stop pulse can vary greatly and such variations are not indicative of the presence of bias. This method of measurement then is applicable not only to circuits using start-stop transmission but also to synchronous transmission systems. Examples of bias measurement for zero, 25-per-cent marking, and 25-per-cent spacing bias are given in Figs. 1(A), (B), and (C), where the duration of the measured single unit signals is indicated by *t*.

In much of the equipment used in the transmission of telegraph signals, bias can generally be removed from the signals by simple adjustments. The other forms of distortion are not so readily corrected and are generally considered as a separate procedure. It therefore seems desirable to measure bias and to measure distortion exclusive of bias as separate quantities. Total distortion then is the summation of these separate quantities. Therefore, for use in this paper, the term distortion will exclude distortion caused by bias. The distortion measurement to be described is based upon the minimum duration of a spacing unit interval. The distortion indication is obtained by observing the particular unit spacing signal whose length differs most from the average time interval and then measuring its departure from the average. An example of the measurement of distortion when the bias is 15-per-cent marking and the distortion other than bias is 25 per cent is shown in Fig. 1(D) in which the short signal indicated for code element 4 is a unit interval of minimum duration. The amount that its length differs from the average is taken as a measure of distortion other than bias.

General Description

The bias and distortion measuring instrument, as designed for general application in field offices and using the operating principles described in the foregoing, is shown in Fig. 2. The instrument is a portable device capable of measuring signal bias and distortion on working telegraph circuits which employ 7 or 7.42 unit code, operating at speeds of 65 or 61 words per minute. Measurement can also be made at 75 or 100 words per minute. The instrument was designed especially for use on 70-milliampere (ma) single-current circuits and polar circuits having signal currents not less than approximately 15 ma. Since the instrument is fundamentally a time-measuring device which indicates the time interval occupied by a unit spacing signal, it can be used for measurement on either start-stop or synchronous systems having signal length and operating currents within the design limits.

The instrument is contained in a metal box 10 inches wide by $10^1/_2$ inches deep by 11 inches high, having a sloping front panel on which two meters are mounted indicating a maximum of 25-per-cent spacing or marking bias and 40-per-cent distortion other than bias. Also located on the front panel are switches for the selection of the operating speed and the measuring circuitry as required and a neon lamp to indicate the reception of spacing and marking signals. Located inside the cabinet are three polar relays, three vacuum tubes, the associated circuit components, and selenium rectifiers operating directly from the 115-volt 60-cycle line to supply positive and negative battery potentials. The polar relays are the standard *202-A* type now generally applied in Western Union service giving reliable and long maintenance-free performance. The circuitry of the instrument is shown in Fig. 3. For descriptive purposes, the circuitry may be divided into four parts, namely the input circuit,

---

Paper 55-166, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing December 13, 1954.

W. D. CANNON is with The Western Union Telegraph Company, New York, N. Y.

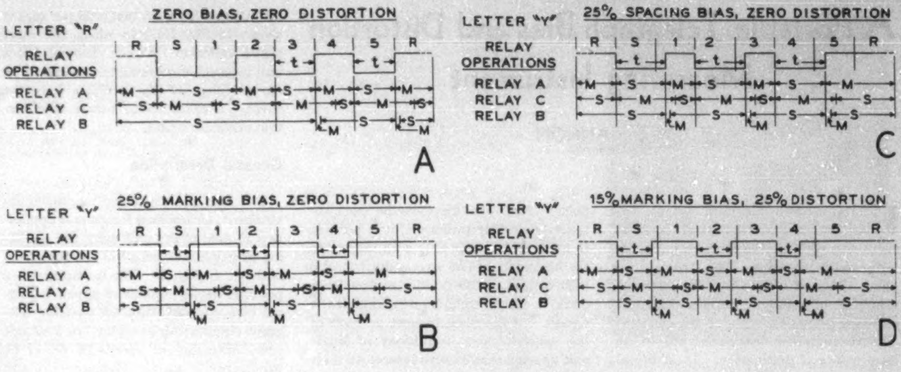


Fig. 1. Theory of bias and distortion measurement

the pulse stretching and transfer circuits, the bias-indicating circuits, and the distortion-indicating circuit.

Input Circuit

The input circuit of the instrument is shown in Fig. 4 in which telegraph signals are impressed upon relay *A*. A relay in the input circuit is desirable since different types of response to the signals are available which may simulate the characteristics of various working equipment. Furthermore, isolation of the input from ground permits the inclusion of a simple power supply for the positive and negative battery potentials.

The operating winding of the relay, having a resistance of 70 ohms and an inductance of 0.45 henry, is inserted directly into the working circuit without appreciably affecting the operating condition of the circuit. Relay *A* follows the signals being transmitted over the circuit, as shown by the flashing of the neon lamp located on the front panel; a spacing signal is indicated when the left half of the lamp is illuminated and a marking signal when the right half is illuminated. A switch is provided to reverse the operating winding of the relay to correspond to the polarity of the signals.

The instrument is arranged to measure either polar signals or single-current signals. For polar signals, relay *A* operates as an unbiased polar relay with the biasing winding open. For single-current signals, it is necessary to include a bias in the relay which is conveniently obtained by a current in a separate winding. This current depends upon the operating current of the transmitted signals and the type of response to these signals which it is desired to incorporate in the instrument. While a number of types of response to the signals are possible, two types, based upon a signal current of 70 ma and designated as *35* and *20-50*, are included in this instrument. Also, these types of response can be applied readily to other signal current values which may vary considerably from the nominal value without appreciably affecting the operation of the instrument.

The 35 and 20–50 Responses

For the *35* type of response, the bias current acting in a direction to space relay *A* is fixed at 35 ma. This relay will then respond to single-current signals in the operating winding and will be operated from the marking contact to the spacing contact from the instant the signal current falls slightly below 35 ma for a spacing signal until the current rises slightly above 35 ma on the following marking signal. Since the transit time and banking current of these relays are small and, also, equal for either a mark-to-space or space-to-mark transition, the elapsed time between the two relay transitions is a measure of the pulse length between 35-ma points on the signal-current wave form. This elapsed time, as interpreted by the *35* type of response, is shown in Fig. 5(A).

For the *20-50* type of response, the bias current in relay *A* has two distinct values prior to each signal transition, depending upon whether the relay tongue is resting against the marking or the spacing contact. This change in biasing current is accomplished by means of tube *T-3*, the grid of which is driven by the potential existing on the tongue of relay *A*. During a marking interval the bias current is 50 ma. Upon the reception of a spacing signal, the current in the operating winding drops from the steady value of 70 ma towards zero, as shown in Fig. 5(B). When the signal current has dropped to 50 ma, the relay is operated from marking to spacing and the potential existing on the relay tongue reverses from positive to negative. This reversal of potential produces a slow transition of the potential on the grid of tube *T-3* and causes the bias current to fall to 20 ma. When a marking signal is received, the signal current rises toward the steady-state value at a rate depending upon the characteristics of the circuit being measured. When this current has risen to 20 ma, relay *A* will be operated from spacing to marking and the resulting reversal of potential will restore the bias current in the relay so that a transition on the next

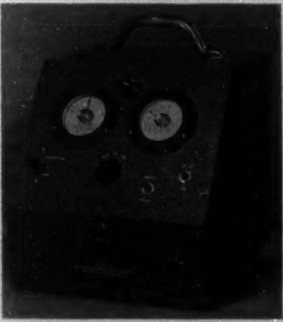


Fig. 2. Front view of field instrument

spacing pulse will take place at approximately 50 ma. Accordingly, a transition in the bias current takes place for each operating transition of relay *A*. This transition in the bias current from one steady state to the other steady state is made gradual by means of the resistance-capacitance network in the grid circuit of the tube so as to prevent false operation on some types of badly distorted wave forms.

Since the transit time and the banking current of relay *A* are equal for either a mark-to-space or a space-to-mark transition, the spacing pulse as measured by the instrument begins when the signal current has dropped to 50 ma and ends when the current has risen to 20 ma. Also, the measuring points on a spacing signal occur at points on the current wave near the beginning of the transition. Hence, the transition delay is usually very small, as shown in Fig. 5(B). Accordingly, the inductive and capacitive characteristics of circuits which may cause a difference between the rise and decay times of the signal current will not be interpreted as bias. Therefore, for all practical purposes, the *20-50* response is unaffected by wave shape and measures the actual time interval of the trans-

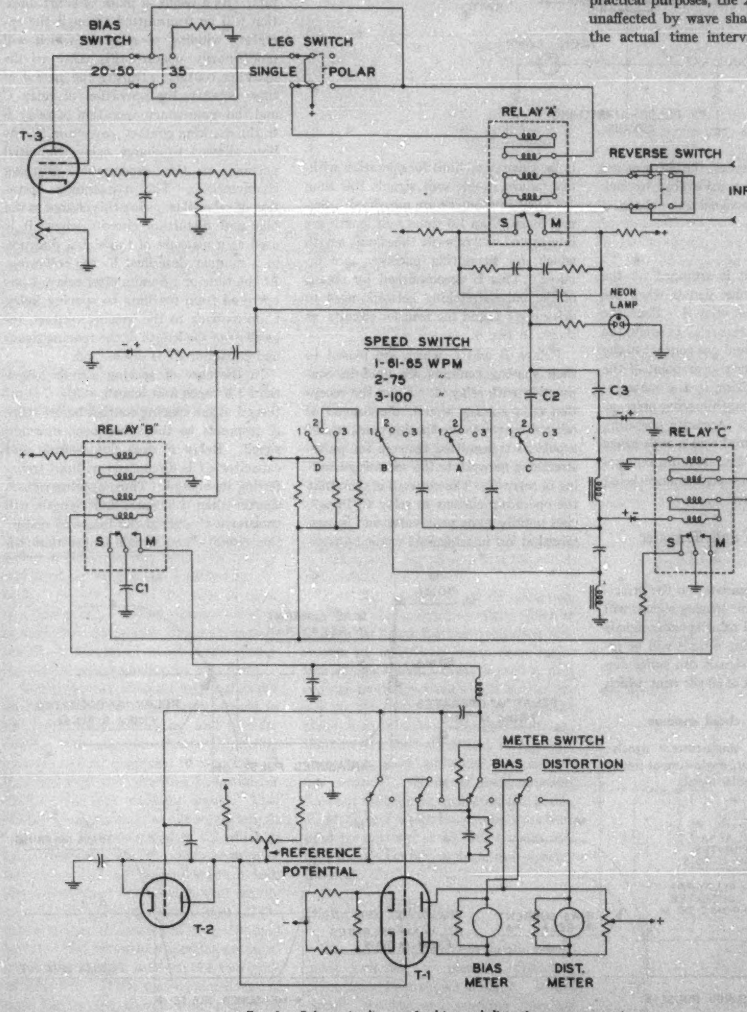

BIAS SWITCH
20-50
35
LEG SWITCH
SINGLE
POLAR
T-3
RELAY "A"
REVERSE SWITCH
INPUT
S
M
NEON LAMP
SPEED SWITCH
1-61-65 WPM
2-75 "
3-100 "
C2
C3
1
2
3
D
B
A
C
RELAY "B"
RELAY "C"
C1
METER SWITCH
BIAS
DISTORTION
REFERENCE POTENTIAL
T-2
T-1
BIAS METER
DIST. METER

Fig. 3. Schematic diagram for bias and distortion measurements

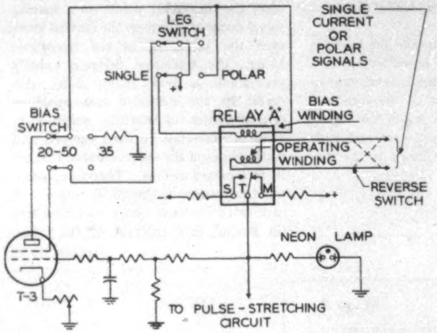


Fig. 4. Input circuit

mitted signals, whereas, the *35* response may be influenced considerably by lack of symmetry in the signal wave shape.

Polar

The input circuit is arranged for the measurement of polar signals when the bias is removed from relay *A*. Since the banking current of this relay is small, the relay will operate from one contact to the other close to the crossover point of the signal current resulting in the measurement of each unit length spacing pulse, as shown in Fig. 5(C). Accurate operating performance is secured over a wide range of signal currents, the minimum requirement for satisfactory performance being of the order of 15 ma.

Pulse Stretching and Transfer Circuits

The instrument operates on the principle that all single unit spacing signals will be measured and all other spacing signals as well as all marking signals will be ignored. Since unit signals can suffer distortion to the extent of 50 per cent which is the theoretical limit for operation without failure, single unit signals less than 1.5 times unit length are measured. Signals longer than 1.5 times unit length are interpreted as 2 or more times unit length which, for measuring purposes are ignored. This is accomplished by means of the pulse-stretching network used to drive relay *C* and the transfer circuits as shown in Fig. 6.

Relays *B* and *C* which are biased to their spacing contacts are used in conjunction with relay *A*. Upon the reception of a spacing signal, the tongue of relay *A* moves to the spacing contact and a pulse is transmitted through the pulse-stretching network to the operating winding of relay *C*. The network is such that the operating current in relay *C*, Fig. 7, rises rapidly to its peak value and is sustained at an intermediate value by resonance after which it rapidly falls to zero. Relay *C* is operated for exactly 1.5 times the duration of a unit length signal immediately following every mark-to-space transition of relay *A* irrespective of the time of occurrence of the subsequent space-to-mark transition. If relay *A* returns to the marking contact after a spacing signal less than 1.5 times unit length, relay *C* will be on its marking contact. As a result a pulse of short duration will be transmitted through the operating winding of relay *B* which will momentarily operate this relay to the marking contact. During the period of time between the operation of relay *C* and the momentary operation of relay *B* to the marking contact capacitor *C1* has been allowed to charge using the initial portion of the exponential charging characteristic. The momentary operation of relay *B* transfers this charge to the bias and distortion circuits where it is used as a measure of the signal duration in a manner described in the following. At the time of 1.5 units after relay *A* has operated from marking to spacing, relay *C* drops back to the spacing contact, regardless of the length of the spacing signal and capacitor *C1* is discharged.

In the case of spacing signals longer than 1.5 times unit length, relay *C* is returned to the spacing contact before relay *A* responds to the subsequent marking signal. Relay *B* does not operate and capacitor *C1* is discharged without transferring its charge. Thus, spacing signals shorter than 1.5 times unit length will transfer a charge to the bias and distortion circuits, and signals longer than 1.5

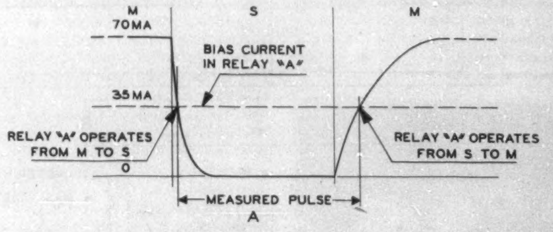
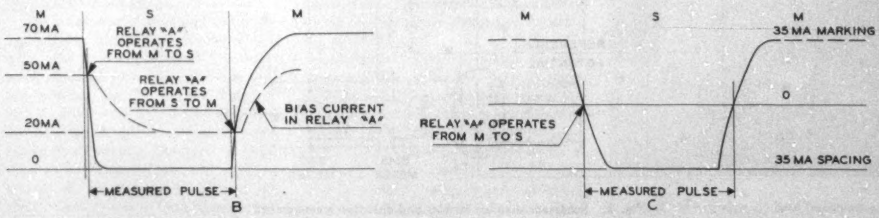


Fig. 5. Input circuit response

A—35 measurement, single-current signals
B—20-50 measurement, single-current signals
C—Measurement of polar signals



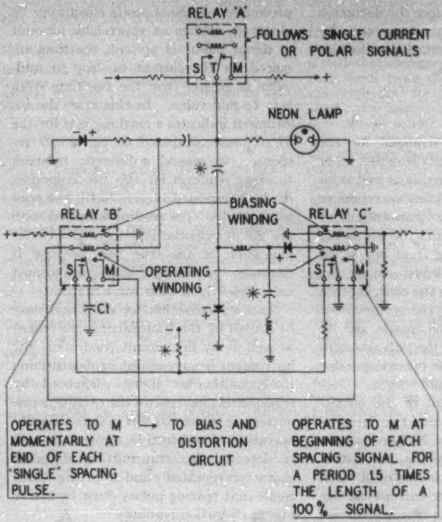

RELAY 'A'
FOLLOWS SINGLE CURRENT OR POLAR SIGNALS
S T M
NEON LAMP
BIASING WINDING
RELAY 'B'
RELAY 'C'
S T M
S T M
OPERATING WINDING
C1
OPERATES TO M MOMENTARILY AT END OF EACH 'SINGLE' SPACING PULSE.
TO BIAS AND DISTORTION CIRCUIT
OPERATES TO M AT BEGINNING OF EACH SPACING SIGNAL FOR A PERIOD 1.5 TIMES THE LENGTH OF A 100% SIGNAL.

Fig. 6 (left). Pulse stretching and transfer circuits

* Values all selected according to operating speed

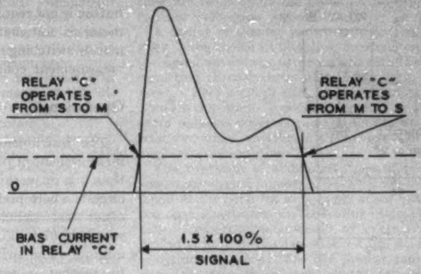


Fig. 7 (above). Operating current in relay C

times unit length will be completely ignored. The sequence of operation of relays *A*, *B*, and *C* is shown in Fig. 1.

Bias Indication

The circuit for bias indication is shown in Fig. 8 in which a twin triode electron tube is arranged in a bridge circuit. The bias meter indicates spacing bias when deflected from the zero center to the left and marking bias when deflected to the right. To one grid of the tube a negative reference potential is applied which is equal to the charge that occurs on capacitor *C1* at the end of a unit length charging period corresponding to zero bias.

When relays *A*, *B*, and *C* respond as described in the foregoing with respect to the input, pulse-stretching, and transfer circuits, the negative charge appearing on capacitor *C1* is applied to capacitor *C2* in series with the reference potential for each single unit spacing signal. This charge, proportionate to the signal length in time, is then averaged in *C2* which is large compared to *C1*. This process of averaging does not depend upon a leak for *C2* but upon the principle that energy may be added or subtracted from *C2* depending upon the length of the individual signal. The potential appearing on capacitor *C2* is applied between the two grids of tube *T-1*. If the spacing signals are of unit length, the charge accumulated on *C2* will be zero and the meter will indicate zero bias. When the spacing signals are shorter than unit length, the potential on *C1* at the time of transfer by relay *B* will be less than the reference potential and *C2* will be charged in such a direction that the meter will indicate marking bias. When the spacing signals are longer than unit length, the potential on *C1* at the time of transfer will be greater than the reference potential and *C2* will be charged in the opposite direction so that the meter will indicate spacing bias.

Since the voltage rise on *C1* occurs on a portion of the charging curve which is reasonably linear the bias indication represents an essentially linear time variation in signal length which is read in percentage deviation from a unit length signal on the bias meter. The indicated bias is independent of the signal combinations transmitted which may be repeated characters, test sentences, or miscellaneous traffic. However, the instrument does not function on characters like blank, *V*, *M*, *O*, or *T* which have sometimes been used for testing, or on other characters that do not contain a single unit spacing interval.

Distortion Indication

The circuit arrangement for the instrument switched to measure distortion other than bias is shown in Fig. 9. In this circuit, spacing and marking bias are ignored and the percentage deviation of any single unit spacing interval which is shorter than the average of the single unit spacing signals is measured with respect to the average. The circuit arrangement is essentially the same as described in the foregoing for indicating bias except that the distortion meter has been substituted for the bias meter; a timing resistor, and a diode rectifier *T-2* have been added and a small capacitor *C3* which was formerly part of *C2* has been switched from parallel to series with respect to *C2*. These components are so arranged that capacitor *C3* instantly recognizes a transfer of potential by relay *B* that is less than the potential stored on *C2* with respect to the reference potential. This occurs whenever a spacing signal shorter than the average has been received. When it occurs, capacitor *C3* is immediately charged to the difference potential through rectifier *T-2* and the distortion meter is deflected up

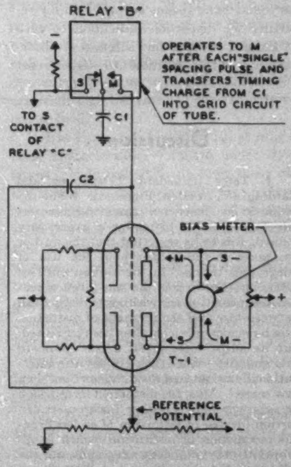

RELAY "B"
OPERATES TO M AFTER EACH "SINGLE" SPACING PULSE AND TRANSFERS TIMING CHARGE FROM C1 INTO GRID CIRCUIT OF TUBE.
S T M
TO S CONTACT OF RELAY "C"
C1
C2
BIAS METER
+M
S−
+S
M−
T-1
REFERENCE POTENTIAL

Fig. 8. Bias-measuring circuit

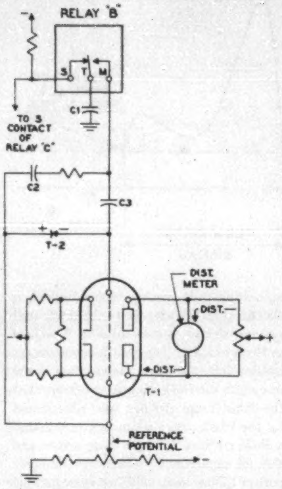


Fig. 9. Distortion-measuring circuit

scale in proportion to the difference. This charge is held on capacitor *C3* by the rectifier so that the meter indicates a steady value. Normally this meter is arranged to drift slowly to zero. If a spacing signal still shorter than any previous signal is received, the meter will be deflected to a new and larger percentage value.

For distortion measurement, it is necessary that some means be provided for resetting the distortion indication to zero at the end of each time interval of observation. In this instrument a special reset button is not required since the distortion meter is automatically set to zero by merely switching momentarily to the bias measurement condition.

Operation

The instrument is arranged for the measurement of bias and distortion other than bias on working circuits as well as on circuits where preferred test sentences or other test signals are transmitted at speeds of 61 to 65, 75, and 100 words per minute. The switching from one speed to another is accomplished simply by varing a resistor to alter the charging rate of capacitor *C1* so that the accumulated charge is the same for all speeds, and by changing a capacitor of the pulse-stretching network to keep the operating pulse of relay *C* at 1.5 times unit length.

Since the input circuit of the instrument is arranged to include two types of response, the *35* and the *20-50* responses, as described in the foregoing, the instrument can be used as an indicator of wave shape symmetry in single-current circuits as well as to measure the bias and distortion on these circuits. For instance, if signals are transmitted over a circuit which possesses the correct proportion of inductance and capacitance, the bias meter will indicate the same bias for either the *35* or the *20-50* responses. However, this condition does not usually prevail. In a circuit having appreciable inductance, and small capacitance, the time required for the current to decay to mid-value is less than the time of rise from zero to mid-value. The instrument indicates this difference as a spacing bias for the *35* response with respect to the *20-50* response. In a circuit where capacitance predominates, the opposite effect may be observed. When an appreciable amount of signal tailing is present, the time required for the current to drop to mid-value is greater than the rise time from zero to mid-value. In this case, the instrument indicates a marking bias for the *35* response compared to the *20-50* response. In general, a difference between the bias readings for the two responses can be interpreted as an indication of poor wave shape. The reading obtained with the *20-50* response is always the better indication of the true bias because it measures the timing of the signals almost independently of their wave shape.

Since excessive bias or distortion may be caused by the transmitting equipment as well as by line circuit conditions the instrument is very useful in determining the condition of signals delivered by transmitting keyboards, distributor transmitters or regenerative repeaters. Faulty operation of a defective segment or cam is detected by transmitting different characters repeatedly and observing which single unit spacing pulses cause excessive bias or distortion readings.

In the routine use of the instrument, measurements are made at both ends of a circuit to evaluate the over-all transmission and detect faulty equipment, poor wave shape, or bias and distortion caused by line-circuit conditions. The accuracy of the instrument in indicating bias percentage is approximately ±1 per cent and slightly less for distortion indication. While this accuracy exceeds the requirements for general application in the field, it is a practice to maintain the calibration of the bias indication so that the readings are never inaccurate by more than 1 per cent.

Vol.
Pt. 1
195

◆

Discussion

V. J. Terry (Standard Telephones and Cables, Ltd., London, England): While too pedantic an insistence upon precise terminology is liable to render a paper very tedious, it is to be regretted that the author has taken so much liberty in the use of telegraphic terms. In the section entitled "Bias and Distortion" he has given a reasoned explanation for wishing to distinguish between bias and the remnant of distortion which remains when bias has been corrected, but he might have given a new name to this quantity rather than employ the internationally recognized word "distortion" in a new sense.[1] Also, the distortion from which he subtracts bias is neither isochronous distortion (degree of distortion of an isochronous modulation or restitution) which is appropriate to synchronous telegraphs, nor the start-stop distortion (degree of distortion of a start-stop modulation or restitution) appropriate to start-stop telegraphs. By considering only the errors in the durations of elements of modulation nominally of unit duration and corresponding to this start condition (space), not only the possibility of worse errors in longer elements of modulation is overlooked but a quantity is measured which is only loosely related to the factors which cause errors of reception in start-stop and synchronous telegraph equipment. The resulting measurements are therefore but a rough indication of the quality of transmission and it would accordingly have been better to have referred to the whole apparatus as a portable "indicator" for telegraph bias and distortion.

If the new device is accepted as an approximate indicator having the virtues of lightness and cheapness, it need not be compared with the precise and relatively complicated distortion measuring sets employed by European administrations, but it is still open to question whether dependence on the adjustment of normal telegraph relays and on an external calibrating device is not too high a price to pay for avoiding a cathode-ray oscilloscope, and for achieving a meter reading which is so well sustained that it leaves no excuse for self-deception on the part of the user. Indeed, a recent oscilloscopic telegraph distortion analyzer of American origin shows how neat and compact a true start-stop distortion-measuring instrument can be made by purely electronic means.

The author has made provision by means of the *20-50* response for assessing the timing of the transmitter contacts with a minimum of error caused by loop inductance, etc. In many cases, however, it is of more interest to assess the suitability of the loop signals to control the magnet of a printer than to determine how perfect the source was from which they originated. In such cases an inversion of the *20-50* response to secure operation to mark at 50 ma and release to space at 20 ma would make the response of the distortion indicator more nearly resemble the average printer. It should be easy to provide this extra facility.

Reference

Documents of the VII Plenary Assembly, *Commité Consultatif International Téléphonique*, Arnhem, Holland, 1953, recommendations *B1* and *B2*, pp. 55-60 of English ed.

W. D. Cannon: One of Mr. Terry's comments on this paper gives consideration to terminology for indicating that remnant of distortion which remains after bias has been eliminated. There is no recognized single term to indicate the summation of the remaining distortion. Therefore, rather than suggest a new name for this quantity, the term "distortion other than bias" was adopted. It would seem that this term in no way conflicts with the generally recognized specific terms for telegraph signal distortion.

In the comments to the effect that the measurements are but a rough indication of signal quality, sufficient information was not included for appraisal of the merits of this contention. A sufficient comparison with other methods of measurement probably would involve too lengthy a discussion for consideration here. However, it can be stated that comparisons have been made with much more elaborate equipment with the result that the performance of this instrument at least meets the requirements for general field service application.

In the comparison of the virtues of a direct-reading device employing relays with an all-electronic cathode-ray oscilloscope instrument, important considerations in favor of the instrument achieving meter readings may be cited as follows. Sustained readings are achieved which are readily discernible both near and at a distance in brilliantly lighted locations. In some applications, an oscilloscope screen requiring close observations and some interpretation is considered impractical. The calibration remains fixed over long periods of time largely owing to the provision of negative feedback and stable relays capable of billions of operations without readjustment. Only relay *A* needs to be accurately but not precisely adjusted and requires checking at much less frequent intervals than the electron tubes. No controls are required on the panel for adjustment or calibration purposes.

A source of greater variation in the results is the choice of sampling positions on the signal wave shape to correspond with the operating characteristics of various equipment. Here a small shift in the sampling position for some types of distorted yet quite operable wave shapes can result in a relatively large variation in the indicated or displayed information.

The provision of additional types of response, of which a number of possible, can be made as the need for such types arises. A response simulating the operating characteristics of a teleprinter would, of course, be desirable for a device intended for use at a teleprinter location. For the present purpose, the two types were selected on the basis that, for all practical purposes, the *20-50* response indicates the true timing of the transmitted signals, and the *35*-type response indicates the performance when a relay is used for reception of the signals. In general, for single-current circuits, a difference in bias in excess of a value of the order of 10 per cent indicates the need for the use of wave-shape corrective measures.

An Accurate Electronic Tracer for Dynamic Characteristics of Magnetic Materials

EMIL KITTL
NONMEMBER AIEE

THIS equipment makes possible the accurate determination of dynamic magnetic characteristics needed for advanced magnetic amplifier design. The hysteresis loop and the time functions of **H**, **B**, and $d\mathbf{B}/dt$ can be either measured point by point or continuously recorded on an *x-y* recorder by means of an automatic electronic a-c compensator which transforms the a-c signal for any given time instant into an equivalent direct voltage. An over-all accuracy of 1 per cent is possible for operating frequencies between 50 and 1,000 cycles per second (cps). The instrument is valuable not only for studying the dynamic characteristics of magnetic materials but also for the evaluation of other nonlinear circuit elements like rectifiers, transistors, and ferroelectric capacitors.

Dynamic Magnetic Core Characteristics and Their Determination

DYNAMIC VERSUS STATIC CHARACTERISTICS

Ballistic methods and static hysteresis loop tracers[1,2] are capable of determining inherent **B H** characteristics. These are, however, greatly modified when the magnetization changes occur at increased speed and/or at greater repetition rates. For these cases the term dynamic magnetization will be used but only the steady-state periodic magnetization processes will be considered in this paper.

The difference between static and dynamic characteristics is mainly caused by eddy currents in the material which, in turn, are governed by the rate-of-flux change, conductivity and thickness of the material, and which produce a broadening of the dynamic as compared to the static hysteresis loop. Additional change in the apparent shape of the dynamic hysteresis loop may be caused by domain wall relaxation, diffusion of impurities, resonance phenomena, and other dynamic effects which are not yet completely understood.

It has long been recognized that magnetic characteristics obtained under dynamic conditions different from those of the actual use are inadequate for quality control and design purposes. The Epstein test was the earliest significant solution to a part of this problem. The comparison of the static hysteresis losses with the dynamic core losses at various frequencies yields valuable information concerning the increase in area of the loop under dynamic conditions. The actual changes in the shape of the loop, however, cannot be determined by this method.

The need for new test methods capable of giving accurate detailed information on the nonlinear characteristics of magnetic materials under dynamic condi-

---

Paper 55-454, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE-APS-AIME-CIT Conference on Magnetism and Magnetic Materials, Pittsburgh, Pa., June 14-16, 1955. Manuscript submitted March 14, 1955; made available for printing May 5, 1955.

Emil Kittl is with the Signal Corps Engineering Laboratories, Fort Monmouth, N. J.

This development work was performed at the request of the Components and Materials Branch, Components Division, of the Signal Corps Engineering Laboratories, Fort Monmouth, N. J. The author wishes to express his gratitude to Dr. E. Both for his excellent support and co-operation in carrying out the rather difficult experimental and theoretical phases of this work. Credit should be given also to W. Schlosser and D. Bange for the careful execution of the design and construction of the development model and the assistance in preparing the diagrams for the report.

Original developmental work on this device is partially covered in the author's doctoral thesis submitted to the Vienna Electrotechnical Institute, Vienna, Austria, in Dec. 1949.

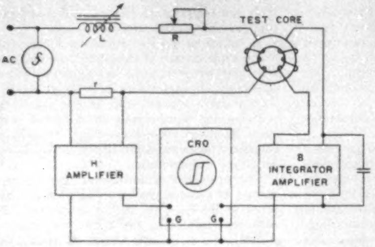


Fig. 1 (left). Oscilloscope loop tracer

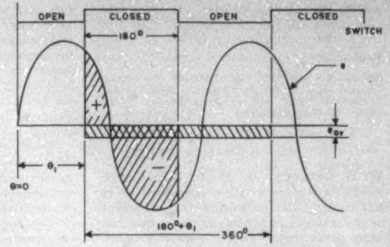


Fig. 2 (right). Basic principle of point-by-point tracing method

tions came about mainly by the rapid development of magnetic amplifiers during the last decade. In this application the magnetic material is subject to wide variations in operating conditions and optimum design can only be accomplished if the performance characteristics of the material for each particular mode of operation is known.

The lack of adequate test equipment has prevented the establishment of core specifications which would fully characterize the attainable core characteristics and has thereby greatly handicapped the progress of this art. While numerous test devices are currently in use which have helped to ascertain at least a minimum performance level for particular designs, it must always be borne in mind that the information obtained from these devices is of limited scope. They do not eliminate the need for test equipment capable of determining all basic core performance characteristics under the whole range of operating conditions prevailing in magnetic amplifiers. The main objective in designing and building the equipment described in this paper was to fill this need as completely as practically possible.

THE QUANTITATIVE DETERMINATION OF DYNAMIC MAGNETIC CHARACTERISTICS

The current in the magnetizing winding of specimen under test is proportional to the magnetizing field. The relation between current and field is given by Ampére's law, taking into account the geometrical configuration of the magnetic circuit

$$\mathbf{H}=\frac{NI^{A}}{l^{cm}} \text{ ampere-turns per centimeter} \qquad (1)$$

If the length of the magnetic path l shows a considerable difference between the shortest and the longest flux lines, a value l_{av} is generally used. However, in this case the apparent shape of the **B-H** loop may be considerably affected by the shape of the core.

The voltage induced in the secondary winding of the specimen is a measure of the rate-of-flux change in this coil. Again certain precautions must be taken to avoid errors caused by the loading effect, leakage flux, distributed winding capacity, etc. The induction law

$$e=N\frac{d\phi}{dt} \quad \text{volts} \qquad (2)$$

$$\phi \ldots \frac{\text{volt-seconds}}{\text{turn}}$$

is the basic relation for this test.

The flux-versus-time function is actually a derived function expressed by the equation

$$\phi(t)=\frac{1}{N}\int e dt = A^{(cm)^2}\mathbf{B}(t)\ \frac{\text{volt-seconds}}{\text{turn cm}^2}\times \frac{\text{volt-seconds}}{\text{turn}} \qquad (3)$$

The dynamic hysteresis loop furnishes meaningful information only if the corresponding voltage-versus-time function is also known. The test equipment should therefore be able to measure and record the following functions:

1. Magnetizing current versus time.
2. Voltage induced in test coil versus time.
3. Flux in test coil versus time.
4. Flux versus magnetizing field (magnetizing current) as a measure of the dynamic **B-H** loop.

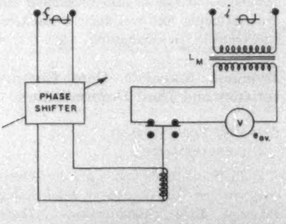


Fig. 3. Circuit for voltage wave analysis using a synchronous switch and a phase shifter

Review of Known Test Methods

OSCILLOSCOPE LOOP TRACERS (FIG. 1)

This method uses a cathode-ray oscilloscope to display the foregoing functions 1 through 4. Several recent publications[3-5] describe this method in detail. Although considerable improvements have been made in the last few years to make commercially available cathode-ray oscilloscopes and associated electronic equipment more reliable, the poor accuracy of evaluation and the inconvenient and laborious method of making durable photographic records place a severe limitation on the use of this method. Its greatest advantage and merit is that it yields simply and quickly semiquantitative information on core material.

POINT-BY-POINT TRACERS USING PHASE-SENSITIVE SWITCHES

The point-by-point tracing method was developed in Germany. Two commercial devices known as the Siemens ferrometer[6] and the A.E.G. vectormeter[7] use a synchronous-driven mechanical contact mechanism in connection with a phase shifter to detect either instantaneous or average values of the applied a-c signal voltages. The basic principle of the point-by-point tracing method[8], Fig. 2,

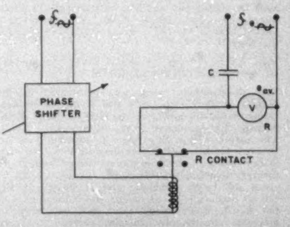


Fig. 4. Circuit for current and flux waveform analysis using a synchronous switch and a phase shifter

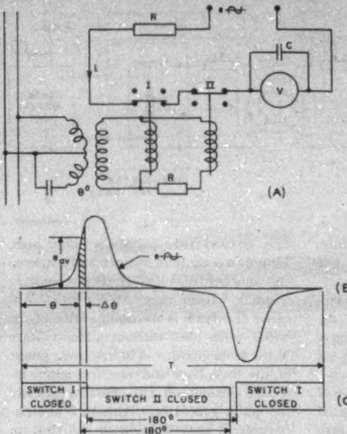


Fig. 5 (left). Circuit using short-time integration to measure instantaneous values by means of synchronous switching devices

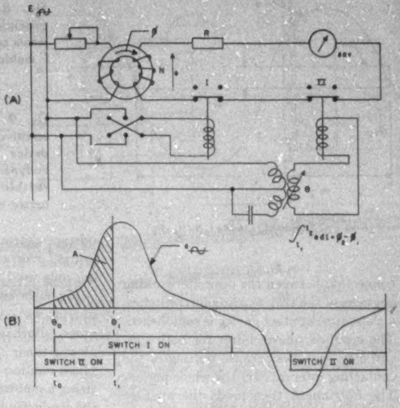


Fig. 6 (right). Circuit performing voltage integration over a variable time interval using two series-connected synchronous switches and one phase shifter

is to assume there is a steady alternating voltage for which by definition the average value over 1 cycle is zero and that the time interval between two instantaneous values, having identical sign and magnitude, is constant and equal to the inverse value of the frequency.

SYNCHRONOUS SWITCHING DEVICES

If the test circuit shown in Fig. 3 is used and it is assumed that the synchronous switch is adjusted to have exactly one half-period of conducting time, then the reading on the voltmeter will be

$$e_{av} \approx \frac{1}{T}\int_{\theta}^{\theta+180°} CR\frac{de}{dt}\,dt = \frac{CR}{T}\,(e_{\theta+180°} - e_\theta) \tag{4}$$

If the voltage e is symmetrical with respect to the zero line and does not contain any even harmonics, then

$$e_{\theta+180°} = -e_\theta \tag{5}$$

and equation 4 is simplified to

$$e_{av} = -\frac{2CR}{T}\,e_\theta \tag{6}$$

In this case, the reading of the voltmeter is proportional to the instantaneous value e_θ and the whole voltage curve can be traced by taking point-by-point readings at properly spaced values of θ between zero and 180 degrees or θ and 180 plus θ.

The disadvantage of this method is that voltages containing even harmonics will give erroneous results because equation 5 is no longer valid. The test method itself introduces a second harmonic, because the circuit impedance is different in successive half-cycles. Therefore, if the impedance of the voltage source is high compared to the external impedance, a considerable error may result. This is particularly true because equation 4 is only valid if

$RC << T$ and $R >> R_{contact}$

The circuit Fig. 4 is mainly used if current or flux curves are to be traced. A mutual air inductance L_M is used for current measurements

$$e_{av} = \frac{1}{T}\int_{\theta}^{\theta+180°} N\frac{d\phi}{dt}\,dt$$

$$= \frac{1}{T}\int_{\theta}^{\theta+180°} L_M\frac{di}{dt}\,dt = \frac{L_M}{T}\times (i_{\theta+180°} - i_\theta) \tag{7}$$

and for symmetrical wave forms

$$i_{\theta+180°} = -i_\theta \tag{8}$$

and

$$e_{av} = -\frac{2L_M}{T}\,i_\theta \tag{9}$$

According to this equation, the average voltage read on the meter is proportional to the instantaneous value of the current for the selected phase angle θ.

The same circuit, Fig. 4, is used for flux measurements except that the mutual inductance is replaced by a suitable 2-winding transformer consisting of the iron-core specimen under test. In this case equation 7 can be modified to

$$e_{av} = \frac{1}{T}\int_{\theta}^{\theta+180°} N\frac{d\phi}{dt}\,dt = \frac{N}{T}\,(\phi_{\theta+180°} - \phi_\theta) \tag{7A}$$

and for symmetrical wave forms in which

$$\phi_{\theta+180°} = -\phi_\theta \tag{8A}$$

$$e_{av} = -\frac{2N}{T}\,\phi_\theta \tag{9A}$$

This principle of testing the flux curve is used in both the Siemens ferrometer and the A.E.G. vectormeter.

A third method, which avoids the restrictions imposed by equations 5 and 8, is shown in Fig. 5. In this case, two synchronous switches are connected in series, both having 180-degree closing time but with a small phase shift between their timing which is introduced by a series resistor in the excitation circuit of one switch only. By making the series

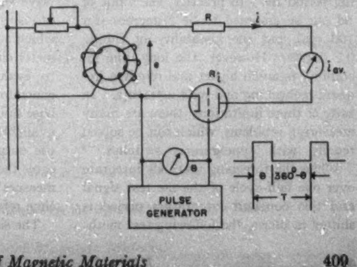


Fig. 7. Circuit to measure instantaneous voltage values by means of an electronic gating device and a d-c meter

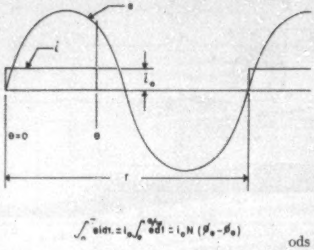


Fig. 8 (left). Basic principle of measurements of phase-sensitive multiplying devices

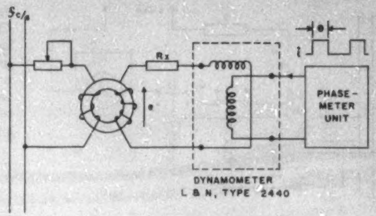


Fig. 9 (right). Phase-sensitive multiplying device using an electrodynamometer and a variable time base square-wave generator

connection between the opposite working contacts of the two synchronous switches, a short-time contact device is established. The common phase shifter for both synchronous switches permits shifting of the contacting interval over the whole period. The instrument then reads the average value

$$e_{av}=\frac{1}{T}\int_\theta^{\theta+\Delta\theta} e\,dt \tag{10}$$

and for small values of $\Delta\theta$

$$\int_\theta^{\theta+\Delta\theta} e\,dt \approx \frac{\Delta\theta}{360}\,T\,(e_{\theta+\Delta\theta/2}) \tag{11}$$

and

$$e_{av}=\frac{\Delta\theta}{360}\;e_{\theta+\Delta\theta/2} \tag{12}$$

This method gives accurate results if $\Delta\theta$ can be made small enough so that $(e_{\theta+\Delta\theta}-e_\theta)/e_\theta$ is within the required accuracy. On the other hand $\Delta\theta$ as a preset parameter of equation 12 must remain constant. Herein lies the difficulty of this method. Synchronous switches, even if built with the highest precision, are always subject to small fluctuations in timing. These are extremely critical in this method and limit the operating frequency to values below 100 cps. Even in this low-frequency range poor performance may be experienced if the tested signal contains high harmonics with large amplitudes for which the resolving power of this method is too small, as it depends on the smallest possible order of the short-time conducting period $\Delta\theta$. In practice, the value of $\Delta\theta$ can be assumed to be 2 degrees (for 100 cps) and the instability of timing 0.5 degree. However, the long-time instability is much higher and requires frequent rechecking and readjustments. In spite of these limitations, there are many measuring problems, which can be solved readily with synchronous switches.[6-8]

While the foregoing methods integrate over one half-cycle of the a-c test signal and this constant integration period is shifted in phase, the following test methods are characterized by a definite integration over a variable time interval and are only used for flux measurements.

Fig. 6 shows a circuit arrangement employing two synchronous switching devices. Both are adjusted for 180-degree closing time. One switch has a constant phase relation in reference to the magnetizing a-c power source while the phase of the conducting interval of the other switch is continuously adjustable between zero and 180 degrees. The adjusted phase angle is read on the calibrated phase-shifting device and the voltmeter reads an average voltage over the effective conducting cycle

$$\int_{t_0=\omega\theta_0}^{t_1=\omega\theta_1} e\,dt = e_{av}T = N\int_{t_0}^{t_1}\frac{d\phi}{dt}\,dt$$
$$= N(\phi_1-\phi_0) \tag{13}$$

The meter reading is proportional to the flux change in reference to the flux value ϕ_0 at t_0.

This method of flux measuring is not limited to symmetrical wave forms and can therefore also be used to measure flux in biased cores. It does not have the disadvantage of the short contact time method described in the foregoing; however, deviations from the adjusted 180-degree closing time and frequency limitations are as severe as in the other methods using synchronous switches.

Electronic Gating and Multiplying Devices

The severe frequency limitation of synchronous switches can be overcome by using electronic gating devices. Fig. 7 shows a simplified circuit. A square-wave signal with the variable time base θ/ω is used to control the grid of a triode, which acts as an electronic switch in the meter circuit. Equation 13 is applicable to evaluate the meter reading. Pulse generators which deliver the variable time base square-wave signal are commercially available. A d-c meter connected across the output of the pulse generator indicates the phase angle directly, which is a measure for the interval of integration or time reference.

The simple circuit shown in Fig. 7 can only be used if the voltage is rather high. Usually a preamplifier is necessary to keep the number of turns in the secondary winding within reasonable limits. Another drawback is the nonlinearity of the triode when used without a bias voltage in the plate circuit. The resolving power of the time base and the frequency response depend on the characteristics of the pulse generator. Although not impossible, it may be pointed out that an over-all accuracy of ±1 per cent affords particular care in the design of electronic equipment.

Multiplying devices can be used instead of gating devices to perform an integration over a variable time interval. In Fig. 8 a square-wave current signal i and a sinusoidal-wave voltage signal e of the same frequency are shown. A meter capable of reading the average product of ei over one complete cycle will show a reading which is proportional to

$$\int_0^T eidt = i_0\int_0^{\theta/\omega} e\,dt = i_0N(\phi_\theta-\phi_0) \tag{14}$$

or

$$W_{av}=\frac{k}{T}\int_0^T eidt = \frac{ki_0N}{T}(\phi\theta-\phi_0) \tag{15}$$

The Electrodynamometer as a Multiplying Device

Fig. 9 shows a circuit diagram in which an electrodynamometer is used as the multiplying device.[9] The phasemeter (or pulse generator) feeds the variable time base square-wave signal i_0 into the field coils F_1 and F_2 of the dynamometer. The moving coils M_1 and M_2 are connected in series with the voltage source e (test coil). The test circuit of Fig. 9 was used successfully in the Signal Corps Engineering Laboratories. A Leeds and Northrup type-*2440* dynamometer was slightly modified for this purpose. With i_0=20-milliampere field current in the parallel-connected coils, the flux sensitivity of this instrument is

$$k=15.10^{-6}\,\frac{R_x}{f}\;\frac{\text{volt-seconds}}{\text{turn}}$$
$$\text{for 1-millimeter deflection} \tag{16}$$

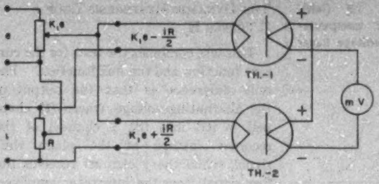
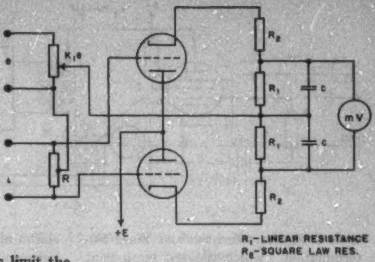


Fig. 10 (left). Phase - sensitive multiplying device using thermocouples

Fig. 11 (right). Phase - sensitive multiplying device using nonlinear resistors



where

R_x = total resistance in voltage path of the dynamometer

f = fundamental test frequency

l = 150 centimeters of effective length of the light beam

Considering a minimum tolerable phase error caused by the inductance of the moving coils, the ratio R_x/f should be kept constant for all frequencies and should not be smaller than $R_x/f \geqq 30$ for a moving-coil inductance of $L = 0.015\mathbf{H}$. Therefore, the flux sensitivity is

$$k = 450.10^{-6} \frac{\text{volt-seconds}}{\text{turn}}$$

for 1-millimeter scale deflection **(16A)**

Example: A core with a cross-section area $A = 1$ centimeter² and a test coil with N = 45 gives a full scale reading of the instrument (250-millimeter deflection) of 2,500 gausses or 10 gausses per millimeter deflection. If this test is made at $f = 400$ cps, $R_x = 30f = 12{,}000$ ohms and the average voltage on the test coil is $E_{av} = 2fN\phi_{max} = 0.09$ volt. The current load is $i_{rms} = 1.11\ (0.09/12{,}000) = 8.25\ 10^{-6}$ ampere and the power losses in the test circuit $i^2R = 0.83\ 10^{-6}$ watt. If, e.g., the sample under test is a ferrite core, magnetized to $\mathbf{B}_{max} = 2{,}500$ gausses, an average figure of its iron losses is about 2.5×10^{-3} watts per gram and even for small samples the power losses in the test circuit are negligible. Tests at 60 and 400 cps have demonstrated an overall accuracy of the flux measurements of ±1 per cent. As the rise time of the square-wave current in the field coils is approximately 3×10^{-6} second, a phase error of 0.5 degree may be expected for a 400-cps fundamental frequency.

Disadvantages: The high sensitivity of the dynamometer requires particular care and maintenance. Special mounting is necessary to eliminate shock and vibration. The minimum time between two readings is at least 10 seconds. A certain zero drift is present and, to obtain the higher accuracy, it is necessary to average two readings taken with different polarity of the potential circuit. The relatively large distributed capacity and the inductance of the dynamometer coils limit the operating frequency to the audio-frequency range. No automatic recording of the flux curve with commercially available recorders is possible. Readings must be taken point by point and the flux function must be plotted manually.

OTHER MULTIPLYING DEVICES

Electronic wattmeters[10,11] are known which are based on elements having a square-law characteristic (current versus voltage). In this case, the instantaneous product of ei is formed as follows: The input circuit consists of a linear network from which the two voltages

$$e_1 = k_1e + k_2i \tag{17}$$

and

$$e_2 = k_1e - k_2i \tag{18}$$

are derived. These two voltages are applied to separate square-law elements in a way which allows the measurement of the average value of the difference

$$e_2^2 - e_1^2 = 4k_1k_2(ei) \tag{19}$$

An instrument reading

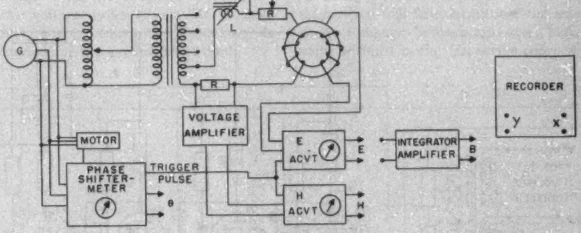
$$e_{av} = \frac{1}{T}\int_0^T (e_1^2 - e_2^2)\,dt = \frac{4k_1k_2}{T}\int_0^T ei\,dt \tag{20}$$

will therefore measure the flux in much the same way as the electrodynamometer.

Two circuits performing the aforementioned task are shown in principle in Figs. 10 and 11. Fig. 10 is the basic circuit for the thermocouple wattmeter and Fig. 11 shows nonlinear resistors as sensing elements. The latter circuit is being used in commercially available wattmeters. Such an instrument, although not particularly suited for the job, was used as a temporary solution to make records of dynamic hysteresis loops on a Leeds and Northrup x-y recorder.

The New Electronic Point-by-Point Tracer

The four different functions, specified in the section entitled "Dynamic Mag-



TO OBTAIN:	TO Y AXIS OF RECORDER CONNECT:	TO X AXIS OF RECORDER CONNECT:
E VS. θ	E OUTPUT OF ACVT	θ
H VS. θ	H OUTPUT OF ACVT	θ
B VS. θ	E OUTPUT OF ACVT THROUGH INTEGRATOR AMPLIFIER	θ
B VS H	E OUTPUT OF ACVT THROUGH INTEGRATOR AMPLIFIER	H OUTPUT OF ACVT

Fig. 12. Block diagram for the dynamic magnetic core tester

ol. 74
. 1
955

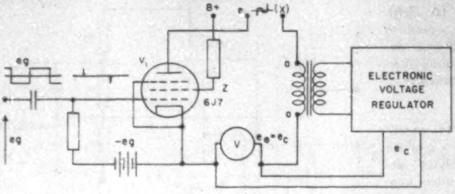


Fig. 13 (left). A-c compensator-voltage tracer

netic Core Characteristics and Their Determination," which the test equipment will perform may be explained by means of the block diagram of Fig. 12.

CURRENT FUNCTION $i=f(\theta)$ OR MAGNETIZING FORCE $\mathbf{H}=f(\theta)$

The block diagram, Fig. 12, shows a conventional magnetizing circuit for the test core. A noninductive shunt *R* produces an input voltage proportional to the magnetizing current which is fed into the **H** amplifier. The output of this amplifier is connected to an electronic alternating-voltage tracer (**H** channel). This unit is controlled by a trigger pulse derived from the phasemeter unit. The trigger pulse selects a particular time instant, for which the alternating-voltage tracer gives the instantaneous value of the input signal wave which can be either read on the output meter or fed into the *y* channel of the *x-y* recorder. The phasemeter unit measures the selected phase angle of the instantaneous current value. This phase angle likewise can be read on a meter or, for automatic recording purposes, a proper signal from the phasemeter unit is fed into the *x* channel of the *x-y* recorder. A small motor-driven phase shifter allows variation of the phase angle between zero and 360 degrees, so that a full cycle of the current (or magnetizing force **H**) can be traced on the recorder.

VOLTAGE FUNCTION $e=f(\theta)$

The operational functions of the units for this test are the same as for current tracing except that the voltage of the secondary winding e_2 is fed directly into the alternating-voltage tracer (channel *E*) As the reference phase angle is not changed, current and voltage traces appear with the correct phase relationship on the recording.

THE FLUX FUNCTION $\phi=f(\theta)$

To trace the flux function, the necessary interconnection between the units, Fig. 12, are altered only in that the output of the *E* channel of the alternating-voltage tracer is fed to an integrator amplifier which delivers the voltage time integral or flux function to the *y* channel of the recorder. Correct voltage time integration in this unit is achieved by a constant and a calibrated time rate of phase-angle change ($d\theta/dt=\lambda$, where λ=a constant) which is provided by a synchronous drive in the phasemeter unit.

THE DYNAMIC HYSTERESIS LOOP ϕ VERSUS H

This test combines the tests for the current function and the flux function. The only difference is that the output of the alternating-voltage tracer (**H** channel) is fed into the *x* channel of the recorder instead of the phase shift signal, while the *y* channel receives the flux signal from the integrator amplifier.

Description of Component-Unit Blocks

A-C COMPENSATOR VOLTAGE TRACER, FIG. 13

The a-c-signal voltage, a variable d-c compensation voltage, a zero-current detector, and a pentode are arranged in a closed-loop circuit. The zero current is detected by means of a pulse transformer which controls the automatic electronic voltage regulator. This unit produces the direct voltage e_c; see Fig. 13. The pentode is normally kept cut off by sufficiently high negative grid bias so that no current can flow in the loop. The square-wave trigger pulse arriving from the phase-shifter unit is differentiated in the *CR* circuit and the resulting short-time pulse across the resistor controls the grid of the pentode. During the duration of the positive pulse (about 2 microseconds), the tube becomes conducting. If at this moment the instantaneous value of the a-c input voltage e_g is less than the d-c compensating voltage, positive plate current will flow and the zero detector will force the electronic regulator to reduce the d-c compensation voltage until equilibrium is reached between the instantaneous value of the a-c input voltage

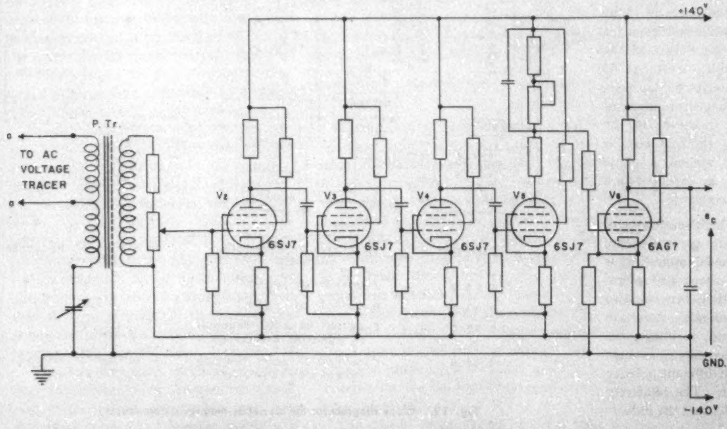
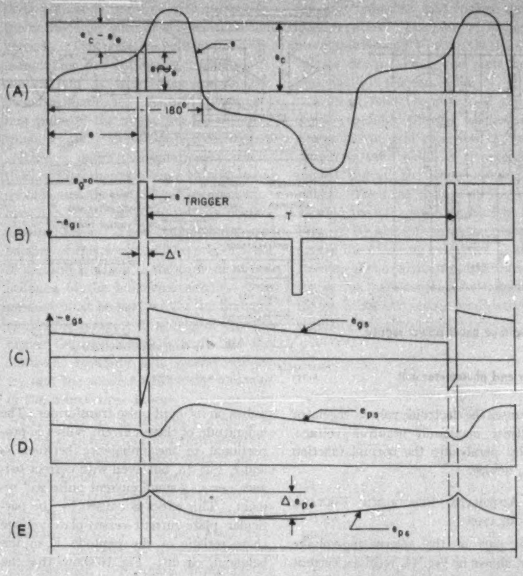


Fig. 14. Circuit diagram, automatic electronic voltage regulator



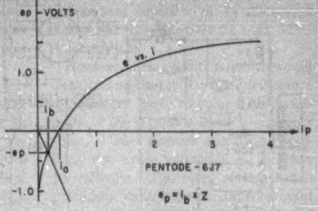


Fig. 15 (left). Wave forms in the automatic voltage regulator

Fig. 16 (above). Plate voltage-current characteristic of the gating tube 6J7

and the d-c compensation voltage. The d-c meter across the output of the electronic regulator will read directly the instantaneous value of the a-c input voltage for the phase angle θ or the time instant at which the positive trigger pulse at the tube occurs. This trigger pulse can be shifted from zero to 360 degrees of phase angle by means of the phase-shifter and phasemeter units and the meter reading of the compensating voltage will exactly trace the a-c input voltage wave form. Thus, the function e versus θ is measured point by point with this instrument.

As the device utilizes a direct compensating measuring method, no calibration is necessary except for the standard multirange d-c voltmeter. The input impedance is almost infinite for balanced condition. Tube noise and instability of the voltage supply limit the voltage sensitivity of the balancing circuit. The sensitivity is about 5 millivolts in the development model and, based on a required accuracy of 0.5 per cent, the lowest full-scale voltage range is 1 volt. Smaller voltages may be measured if accuracy is sacrificed or if a preamplier is used.

High-frequency transients which can enter the circuit at the input side may cause slow transients and noise in the output. Therefore, shielding of the a-c signal source and the connecting input leads are of particular importance if high-impedance sources are used. A zero-adjustment control is provided in the electronic voltage regulator to adjust zero output when the input terminals are short-circuited. The output voltage may be affected if the signal source has extremely high impedance. The small sensing-current pulse in the compensating circuit can cause considerable voltage drop in the source impedance which will increase the voltage across the input terminals and cause a positive error in the output voltage. If the signal-source impedance is constant, this error may be compensated by the zero-setting control. However, in the case of a high-impedance test coil on a magnetic core sample, which will be magnetized well into saturation, the impedance changes between zero and a high value within 1 cycle. In severe cases, it

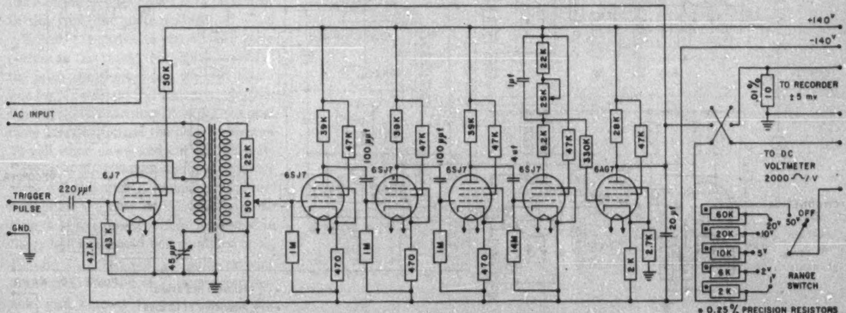

AC INPUT
TRIGGER PULSE
GND.
220 μμf
6J7
50K
47K
43K
45 μμf
22K
50K
6SJ7
39K
47K
1M
470
100 μμf
6SJ7
39K
47K
1M
470
100 μμf
6SJ7
39K
47K
1M
470
4 μf
6SJ7
1 μf
22K
25K
8.2K
47K
330K
14M
6AG7
28K
47K
2K
2.7K
20 μf
+140V
-140V
.01 %
10
TO RECORDER ±5 mv
TO DC VOLTMETER 2000 Ω/V
60K
20K
10K
6K
2K
50V
20V
10V
5V
2V
1V
OFF
RANGE SWITCH
* 0.25% PRECISION RESISTORS

Fig. 17. Detailed circuit diagram for the alternating-voltage tracer unit

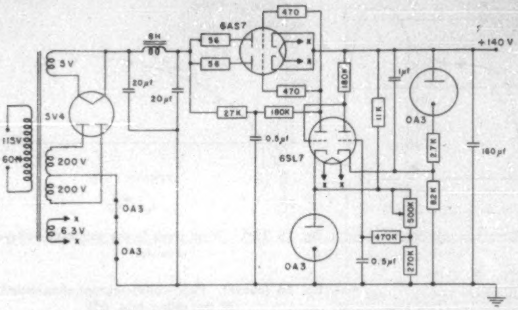

470
6AS7
8H
56
56
470
+140 V
5V
20μf
20μf
5V4
115V
60~
200 V
200 V
6.3 V
x
OA3
OA3
27K
180K
0.5μf
180K
11K
1μf
OA3
2.7K
160μf
6SL7
500K
82K
470K
OA3
0.5μf
270K

Fig. 18 (above). Circuit diagram for the alternating-voltage tracer power supply

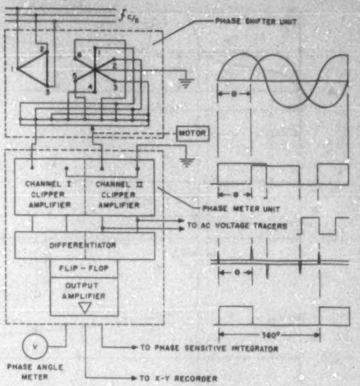

f c/s
PHASE SHIFTER UNIT
1
2
3
4
5
6
MOTOR
CHANNEL I CLIPPER AMPLIFIER
CHANNEL II CLIPPER AMPLIFIER
PHASE METER UNIT
TO AC VOLTAGE TRACERS
DIFFERENTIATOR
FLIP - FLOP
OUTPUT AMPLIFIER
θ
360°
TO PHASE SENSITIVE INTEGRATOR
V
PHASE ANGLE METER
TO X-Y RECORDER

Fig. 19 (right). Block diagram for phase shifter and phasemeter unit

may be necessary to use a cathode-follower input stage.

For the present laboratory model of this unit, the upper frequency limit is about 10,000 cps. It is mainly determined by the capacitive current of the tube (plate-to-cathode capacity) which in this frequency range reaches the magnitude of the sensing current of the zero detector and causes the electronic voltage regulator to adjust maximum negative voltage, thereby paralyzing the normal function of the device.

The Automatic Electronic Voltage Regulator

This part of the alternating-voltage tracer, shown in Fig. 14, receives current pulses on its input pulse transformer. The magnitude of these current pulses is proportional to the unbalance between e_θ and e_c Fig. 15, but even with correct balance $e_\theta = e_c$, a small current pulse will remain. This effect is caused by the particular plate-current-versus-plate-voltage characteristic of the pentode V in the balancing circuit. Fig. 16 shows that the

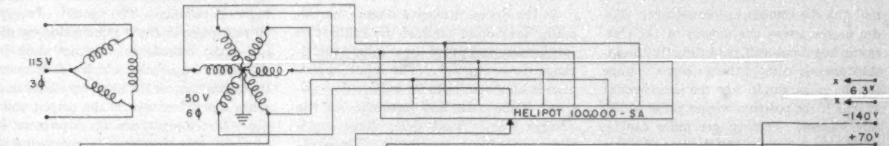

115 V
3φ
50 V
6φ
HELIPOT 100,000 - SA
6.3V
-140V
+70V
+310V
PHASE ANGLE METER
2000Ω 1V F.S.
80μf
11246
OFF
360°
180°
90°
36°
288
304
658
249.9
224920
112460
67475
42985
CHANNEL I
CHANNEL II
7F7
7N7
6SN7
6AG7
220K
56K
100K
120K
56K
.05μf
22K
26K
26K
100K
68K
470
1M
5-50μμf
9K
9K
500K
174K
50K
100μμf
80K
IN34
11K
56K
1K
750
20
50
RECORDER
23.1
23.1
46.2
138.6
PULSE TO ACVT
• 1% PRECISION RESISTORS
• 0.1% PRECISION RESISTORS

Fig. 20. Detailed circuit diagram for the phase shifter and phasemeter unit

plate voltage is negative with respect to the cathode for current smaller than the value i_0. This negative voltage can compensate for the voltage drop caused by the current i_b which remains in the balanced condition. With the input current pulse of the magnitude i_b, the electronic regulator should deliver the output voltage e_c under quasi-stationary conditions. That means a small change $\pm\Delta i_b$ should cause the output voltage to change over the whole range of voltage control required for e_c. Then the maximum unbalance of the circuit will be proportional to Δi_b and z, which is the equivalent impedance of the balancing circuit. This product must be held within the limits of required accuracy. It is evident that the higher the gain which is built into the electronic regulator, the smaller will be Δi_b and the smaller will be the influence of the source impedance.

Connected to the secondary winding of the pulse transformer is a resistance network consisting of a fixed resistor and a potentiometer which allows the adjustment of the grid input to the following 3-stage a-c amplifier, comprising three *6SJ7* pentodes, V_2 to V_4. The next stage, pentode V_5, operates as a peak indicating grid rectifier. The plate output of this tube is essentially a direct voltage which is proportional to the peak value of the input signal at the capacitor. The peak rectifier characteristic is accomplished by using a very high time constant $T=CR$ in the circuit which couples to the grid. Resistance coupling is used to the tube V_6(*6AG7*) which regulates the output voltage e_c.

To understand the functional operation of this electronic regulator assume that the output voltage e_c of the electronic regulator, which is equal to the difference between the plate voltage e_p of V_6 and ground, is smaller than the instantaneous value of the signal voltage e_θ to be tested. In this case, the plate voltage of V_1 is sufficiently negative to cut off any plate current in this tube. No signal arrives at the pulse transformer and the grid capacitor for V_5 will slowly discharge through the high grid resistance, making the grid more positive so that the plate voltage of V_5 will come down and cut off the plate current in V_6. Hence, the plate voltage of V_6 will rise until $e_c>e_\theta$ and a plate current pulse will flow in V_1. At the same time, a voltage pulse will be generated in the pulse transformer and amplified in V_2, V_3, and V_4. The positive pulse arriving at the grid capacitor of V_5 causes considerable grid current to pass through the capacitor thereby charging it. After the pulse is over, a higher voltage appears at the grid capacitor and makes the grid more negative. The plate of V_5 becomes more positive and V_6 controls the output voltage down. Thus, the plate current pulse in V_1 is reduced again.

A steady-state condition of the electronic regulator will be reached if the charge of the grid capacitor of tube V_5 caused by the pulse is completely removed during the time between two pulses. Then the voltage on the grid of V_5 will have a shape as shown in Fig. 15(C). Figs. 15(E) and (D) show the plate voltages of the following two tubes V_5 and V_6. Because of the particular conditions existing in this pulse-controlled electronic regulator, other stationary conditions of the regulator will theoretically exist for plate-current pulse magnitudes for which the grid current capacitor of V_5 discharges within the time intervals $2T-\Delta t$, $3T-\Delta t$, or $nT-\Delta t$.

However, these conditions which would mean a considerably increased unbalance $e_c<e_\theta$ are not likely to occur if the sensitivity of the regulator is set to a proper value. Under normal operating conditions, the ripple voltage of e_{p5} is below 10 millivolts. The speed of response is practically limited by the time constant in the grid circuit of V_5 to a maximum of a 10-volt-per-second change of e_θ. This in turn will determine the proper rate of phase shift of the trigger signal. For instance, to trace a sinusoidal voltage signal of 50 volts' peak value, the phase-shifter

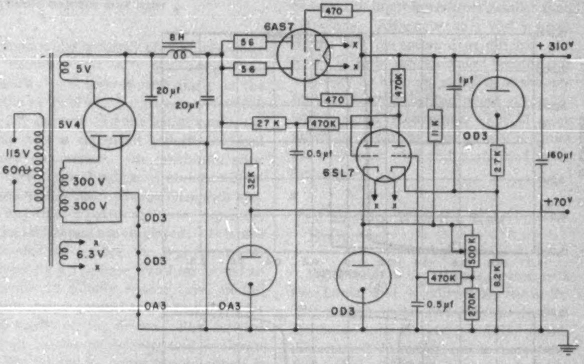

470
6AS7
8H
56
56
+310V
5V
1μf
20μf
20μf
470
470K
5V4
27K
470K
11K
OD3
0.5μf
160μf
115V
60~
6SL7
27K
32K
300V
300V
+70V
OD3
500K
6.3V
OD3
8.2K
470K
OA3
OA3
OD3
0.5μf
270K

Fig. 21. Power supply circuit for phasemeter unit

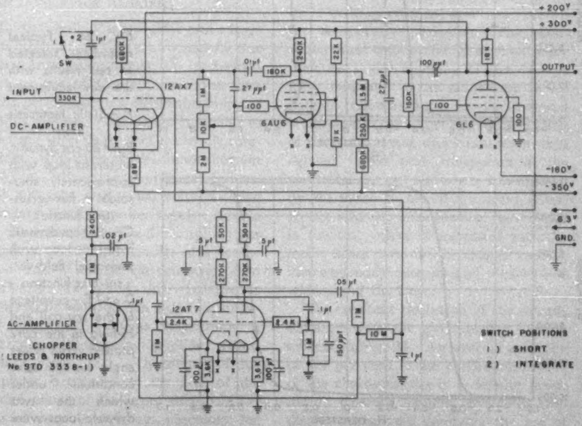

+200V
+300V
2
1μf
SW
680K
240K
22K
18K
.01μf
180K
100μμf
OUTPUT
INPUT
330K
12AX7
1M
27μμf
100
15M
27μμf
150K
100
DC-AMPLIFIER
10K
6AU6
250K
6L6
100
2M
1.8M
11K
680K
-160V
-350V
6.3V
GND.
240K
.02μf
50K
50K
5μf
5μf
1M
270K
270K
.05μf
.1μf
12AT7
24K
24K
.1μf
1M
AC-AMPLIFIER
CHOPPER
LEEDS & NORTHRUP
No STD 3338-1)
1M
1M
150μμf
10M
1μf
3.6K
100μf
3.6K
100μf
SWITCH POSITIONS
1) SHORT
2) INTEGRATE

Fig. 22. Detailed circuit diagram for the integrator amplifier

ol. 74
. 1
55

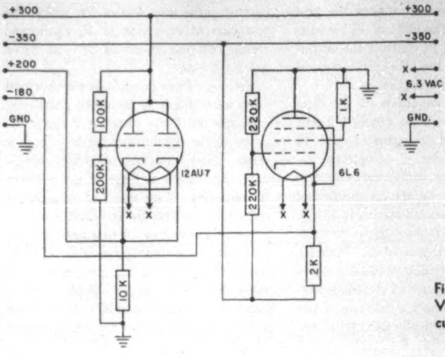


Fig. 23 (left). Voltage divider circuits for the integrator amplifier

speed should be adjusted so that it runs through 1 complete cycle in not less than $T=2\pi e_{\text{peak}}/10=31.4$ seconds.

The sensitivity of the electronic regulator can be varied in the grid circuit of tube V_2. An increase in sensitivity will have the effect that a smaller pulse current is required to establish the stationary condition. As can be seen from Fig. 16, a smaller plate current in V_1 means a larger negative plate voltage in V_1 and a smaller voltage drop in the balancing circuit. The result is that with a given signal-source impedance the indicated output voltage will be reduced if the sensitivity of the electronic regulator is increased. As already mentioned, this effect is used for zero-level adjustment at a given source impedance.

A separate d-c power supply with electronically regulated outputs for +140 volts and −140 volts is employed for the automatic electronic voltage regulator. The ripple in the output voltages should be kept smaller than 10 millivolts if a square-wave current load of 20 milliamperes is applied. The detailed circuit diagram of the laboratory model alter-

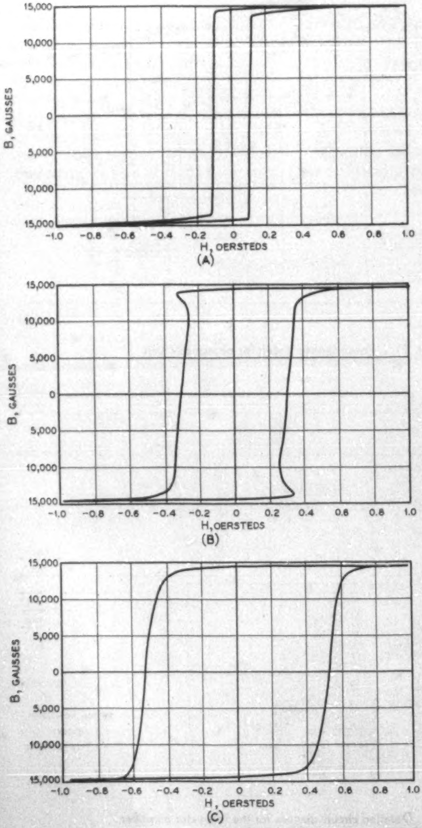


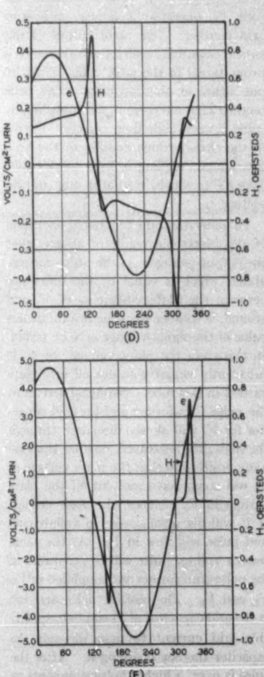
Fig. 24. Typical recordings selected as test results with the new equipment

A—Static hysteresis loop

B—420-cps dynamic hysteresis loop with approximately sinusoidal flux-versus-time function

C—420-cps dynamic hysteresis loop with sinusoidal field-versus-time function

D and E—e_2-voltage in the test coil and H-field in the core plotted versus phase angle for the same conditions under which the two dynamic loops were plotted. f = 420 cps



Vol. 7
Pt. 1
1955

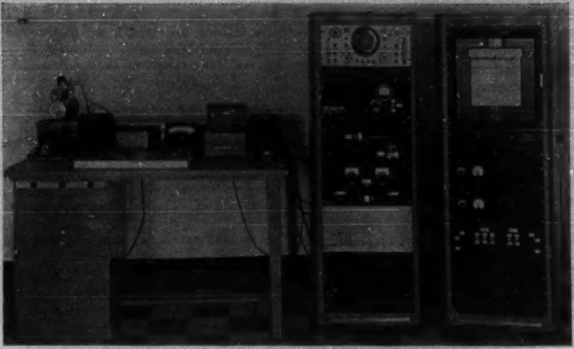


Fig. 25. Laboratory model of the dynamic magnetic core tester

nating-voltage tracer is shown in Fig. 17 and the two identical 140-volt power supplies are in Fig. 18.

PHASE-SHIFTER AND PHASEMETER UNIT

A small 3-phase tapped-potentiometer type of phase shifter is used; see Fig. 19. A 3-phase to 6-phase transformer minimizes the amplitude changes of the phase-shifted voltage. The 10-turn helically wound potentiometer has nine taps and allows a 540-degree phase shift with excellent sensitivity of phase-angle adjustment. A small geared motor drives the potentiometer with constant calibrated speed if continuous recording is required. A reference voltage is taken from the center point of the secondary transformer winding and from the starting point of the potentiometer which indicates zero phase angle. This reference voltage and the phase-shifted voltage are connected to separate channels of the phasemeter unit. First, the sinusoidal signals are converted into sharp rising square-wave signals in a clipper amplifier. These square-wave signals are then differentiated to form sharp periodic pulses which are spaced in accordance with the phase angle. Each of the periodic pulse signals operates one side of the Eccles-Jordan flip-flop circuit which gives a single square-wave plate output of constant amplitude and a time base varying in accordance with the phase angle θ. Three different signals can be obtained from this phase-angle meter. The phase angle θ can be read directly on an accurate (0.25-per-cent) d-c meter with 360-, 180-, 90-, and 36-degrees full-scale readings; a 10-millivolt signal is used for the x-y recorder, and a 20-milliampere current signal can be used for the phase-sensitive integrator. At low frequencies (below 100 cps), the performance and accuracy of the phasemeter is excellent. At 400 cps, some nonlinearity is observed particularly in the initial range of 0 to 5 degrees and from 355 to 360 degrees. In the initial range, the time axis is expanded and in the final range compressed; the maximum error in these critical ranges is about ±2 degrees at 400 cps. The performance may be improved if better and more expensive pulse techniques are employed. It is also intended to replace the 6-phase transformer by a 12-phase transformer and to provide 12 equally spaced taps on the phase-shifter potentiometer to reduce the unlinearity of this device. The detailed circuit for the phasemeter unit is shown in Fig. 20 and its power supply circuit is shown in Fig. 21.

THE INTEGRATOR AMPLIFIER

Exceptionally low zero drift and a large time constant are the requirement for this unit. Integration must be performed over a period of 10 minutes with less than ±0.5-per-cent error of the maximum rated output voltage. The design is similar to the type which is used in electronic computers. Fig. 22 shows the circuit diagram and Fig. 23 the power supply connections. For 300 volts and −350 volts d-c two standard lambda-power-supply models no. *28* were used. The d-c amplifier has a gain of 25,000 and consists of a double triode *12AX7* and the two pentodes *6AU6* and *6L6*. Improved drift balance is achieved by addition of a drift-free amplifier comprising a 60-cps chopper and a conventional 2-stage a-c amplifier with a gain of about 1,000. The output of this amplifier is coupled to the grid of the second triode of the input tube *12AX7*. Special requirements are necessary for the feedback capacitor, which must have a leakage resistance higher than 2,000 megohms at maximum output voltage (50 volts) of the integrator. High thermal stability of the rated capacity value is also required. Performance and long-time stability of a breadboard model of this integrator were not satisfactory. At present, the unit is being rebuilt and will be improved.

THE X-Y RECORDER

A standard Leeds and Northrup x-y recorder, model Speedomax, type *G*, has been used successfully with the equipment. It has a sensitivity of 1 millivolt per inch, 10 inches of maximum deflection in both axes, and a maximum speed of 2.5 inches per second. To utilize the full accuracy of this instrument, it is advisable to keep the input impedance smaller than 100 ohms. Although the x and y input signals can be operated at different d-c levels, it is not recommended because it is necessary to keep the stability of this potential difference within about 0.1 millivolt to avoid interference. The Moseley autograf, a similar type of x-y recorder has also been used and performed satisfactorily.

Test Results With the New Equipment

Fig. 24 shows the records plotted with the Leeds and Northrup x-y recorder. The sample under test was an Arnold-Deltamax toroidal core no. *5387*, outside diameter = 1.75 inches, inside diameter = 1.25 inches with 0.002 inch of tape thickness. The static loop was traced with a static loop tracer,[3] using the same Leed and Northrup recorder, and is only shown for comparative purposes. The sinusoidal flux test is the most commonly used test; the widening of the 420-cps loop compared to the static loop is apparent. Characteristic for the material is the peak in the **H** curve at the beginning of the flux reversal, which is at about 0.32 oersted. This peak disappears in the sinusoidal current test where it is masked by the much higher eddy-current effect. Notice also the pulse form of the voltage in this test, which is correlated with the much faster rate-of-flux change causing increased eddy currents and considerable widening of the loop.

The excellent resolution of the records obtained with the instrument is clearly evident in Fig. 24. An over-all accuracy of ±1 per cent in the evaluation of any of the characteristics is a definite possibility, whereas present hysteresis-loop display techniques involving the use of CRO tubes are at best capable of accuracies in the order of ±3 to ±5 per cent.

Conclusions

The electronic tracer is a new and valuable tool for the accurate and convenient determination of magnetic characteristics under dynamic condition. It will, therefore, offer great assistance in the development of magnetic materials as well as in the engineering and production of magnetic components. Dynamic performance characteristics of semiconductors, transistors, rectifiers, and other nonlinear components may also be tested under improved conditions. Therefore, the instrument really fills a long-felt need between the slow response recording oscillographs and the high-frequency CRO.

References

1. A Recording Fluxmeter of High Accuracy and Sensitivity, P. P. Cioffi. *Review of Scientific Instruments*, New York, N. Y., vol. 21, July 1950, pp. 624–28.

2. Electronic Integrating Fluxmeter. *Research Report*, Philips Laboratories, Inc., June 1953.

3. Dynamic Hysteresis Loop Measuring Equipment, H. W. Lord. *Electrical Engineering*, vol. 71, no. 6, June 1952, pp. 518–21.

4. Dynamic Hysteresis Loops of Several Core Materials Employed in Magnetic Amplifiers, H. W. Lord. *AIEE Transactions*, vol. 72, pt. I, March 1953, pp. 85–88.

5. Evaluation of Core Material for Magnetic-Amplifiers, D. C. Dieterle. *Electrical Manufacturing*, New York, N. Y., Jan. and Feb. 1953.

6. The Siemens Ferrometer (in German), Dr. Thal. *Archiv fuer Technisches Messen*, Munich, Germany, J-60–243.

7. Method of Measuring Rectangular Hysteresis Loops of Toroidal Cores, F. Koppelmann. *Archif fuer Elektrotechnik*, Berlin-Charlottenburg, Germany, vol. 38, nos. 3 and 4, 1944, pp. 141–56.

8. *Archiv fuer Technisches Messen*, Munich, Germany, V362–1, V951–2, F540–1,2.

9. Emil Kittl. *Austrian Patent No. 166.027*, Vienna, Austria, Aug. 10, 1950.

10. Thermo Couple Audio Frequency Meter, J. B. Ryder. *Proceedings*, National Electronics Conference, Chicago, Ill., 1952.

11. A Stabilized Electronic Multiplier, C. D. Morrel, R. V. Baum. *Ibid*.

Temperature Errors in a Dragmagnet Eddy-Current Disk Type of Tachometer Indicator

L. T. AKELEY
MEMBER AIEE

J. J. FRAIZER
NONMEMBER AIEE

Synopsis: An analysis has been made of the factors producing temperature error in a dragmagnet eddy-current disk type of tachometer indicator. A new design for tachometer compensation is described and test results are presented.

TODAY'S military aircraft are required to operate under conditions where large variations in temperature are encountered. Despite these variations, aircraft instruments are required to meet very high standards of performance and accuracy.

Aircraft tachometer indicators, like most instruments, are subject to change in indication with change in temperature. This is called temperature error. Tachometer indicators used in high-speed military aircraft will be subject to large and sometimes rapid temperature changes since these aircraft must operate in any region of the earth and, in many cases, must go from low to high altitudes very quickly.

Most of today's aircraft use turbojet engines whose starting and operating speed must be accurately known. This need for accurate speed indication over a large range of ambient temperature reduces the acceptable temperature error of aircraft tachometer indicators to approximately 1/2 per cent.

The purpose of this paper is to provide a better understanding of the source of temperature errors in the dragmagnet eddy-current disk type of tachometer indicator. To accomplish this purpose the following is done:

1. The source and magnitude of all temperature errors in the indicator are determined.
2. An improved temperature-compensated dragmagnet design is described and its performance compared with that of an earlier design.

Physical Concepts in the Tachometer System

Fig. 1 shows schematically the basic elements in the aircraft tachometer system. The permanent magnet rotor of an electric generator is driven by the source of rotation to be measured. The voltage and frequency so generated is fed to a synchronous motor which drives the dragmagnet assembly at a speed which is an integral function of the unknown revolutions per minute. The dragmagnet assembly consists of 12 opposed permanent magnets with an air gap in which the eddy-current disk is free to rotate. The principal flux path is through this air gap. The nickel-steel shunt serves to divert an amount of this flux which is dependent on the temperature.

The rotating flux in the air gap of the dragmagnet assembly, the permanent magnet poles of which are alternately north and south, produces currents in the eddy-current disk; these react with the dragmagnet field to produce a torque which tends to rotate the disk in the direction of the rotating dragmagnet assembly. The eddy-current disk shaft, restrained by a torque spring, rotates until the spring torque equals the torque from the dragmagnet, the angular position of the shaft and the indicating pointer attached to the shaft then being a measure of disk torque. Since the disk torque is proportional to dragmagnet angular velocity, it follows that the angular position of the pointer may be used to indicate generator revolutions per minute, the quantity to be measured. If the torque spring is

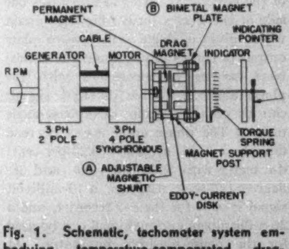


Fig. 1. Schematic, tachometer system embodying temperature-compensated dragmagnet torque elements

---

Paper 55-242, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 20, 1954; made available for printing March 1, 1955.

L. T. Akeley and J. J. Fraizer are with the General Electric Company, West Lynn, Mass.

Table I. Contribution of Particular Instrument Elements to the Temperature Error of Uncompensated Tachometer Indicators

Part	% Temp. Error Per 100 C Average Value	% Temp. Error Per 100 C Expected Tolerance
Group A		
Eddy-current disk, aluminum-manganese alloy	−6.0	±2.0
Dragmagnet, Alnico V	−5.0	±1.0
Air gap (brass posts)	−1.0	±0.1
Torque spring, phosphor bronze	+3.8	±0.2
Total	−8.2	±3.3
Group B		
Eddy-current disk, copper-manganese alloy	−6.0	±1.0
Dragmagnet, Alnico V	−5.0	±1.0
Air gap, invar posts	+1.0	±0.1
Torque spring, phosphor bronze	+3.8	±0.2
Total	−6.2	±2.3

linear, the tachometer indicator dial may be linear and, with the proper gear train between the disk and the pointer, the scale may be made as long as desired for easy reading.

Unfortunately, even a system as basically simple as the foregoing is subject to error because of the physical limitation of its parts. Particular attention will be given to those tachometer indicator errors which accompany change in temperature.

Table I lists known sources of temperature error in the tachometer before compensation. These are change in resistance of the eddy-current disk, change in strength of the dragmagnets, change in dragmagnet air gap because of relative expansion of posts and magnets of the dragmagnet assembly, and change in stiffness of the indicator spring. All of these factors contribute to a change in tachometer indication with temperature change. Together these sources of temperature error should add to a net magnitude of error in agreement with the test results of Table II, column 2, and with the average uncompensated indicator error curve of Fig. 2. When comparing these results it must be remembered that the indicator represented by curve *A*, Fig. 2, has invar posts; the indicators of Table II, column 2, have brass posts and hence a 2-per-cent (%) larger error.

In Table I a value and a probable tolerance variation has been estimated for each of the four sources of temperature error. The effect of temperature upon the resistance, expansion, and stiffness of most metals and metal alloys is well known. Somewhat less is known about the change of magnet strength with temperature. As indicated in Table I, the net error of the uncompensated indicator has a negative sign, i.e., the indicator reading becomes lower with increase in temperature. This net error will obviously be reduced by any change which reduces the magnitude of the negative error elements or which increases the magnitude of the only positive error element, the torque spring.

The temperature coefficient of resistance of the eddy-current disk material can be reduced but only at the expense of considerably decreased disk torque and resultant increased indicator friction error. This, of course, is undesirable.

The magnet material used in the dragmagnet assembly is Alnico V, which is one of the higher energy magnet materials. Characteristically, these magnets provide a large amount of flux at low temperatures and a lesser amount at high temperatures. This change results in one of the fundamental errors of the system since the eddy-current disk torque is a function of the dragmagnet field squared. Because of this relationship, the temperature error caused by a change in magnet strength is approximately double the change of magnet strength for a given temperature change.

The change in field strength in the air gap caused by greater expansion of the post than of the magnet can be reduced by the use of invar in place of brass for the magnet assembly posts. This use of invar for posts is practical and, as indicated in group *B* of Table I, reverses the sign of the air-gap error.

Table II. Temperature Error Tests on Five Uncompensated Tachometer Indicators, Random Samples With Brass Posts, Aluminum-Manganese Disks

Sample	% Error −55 C to +70 C	% Error −30 C to +70 C
1	−12.9	−10.2
2	−13.5	−10.9
3	−10.3	− 9.3
4	−11.0	− 8.8
5	−11.0	− 8.5
Average	−11.7	− 9.5

Compensation Means and Calculations

The tachometer indicators of group *B*, Table I, were designed to reduce the inherent error before applying compensation as well as to provide more effective compensation. The tachometers of group *A* are the previous uncompensated design. The design differences between groups *B* and *A* include:

1. The supporting posts in group *B* are made of invar whereas those of group *A* are made of brass. As is noted in Table I, this results in a change of sign in the temperature error because of the change in relative expansions of posts and magnets.
2. The eddy-current disks in group *B* are made of copper manganese while group *A* uses aluminum manganese. Although this does not result in a direct reduction in temperature error, it does result in a lower tolerance because of the fact that large ingots of the copper-manganese material can be cast with more uniform temperature coefficients.
3. The material of the upper magnet plate has been changed from steel in group *A* to a nickel-invar bimetal[1] in group *B*.

The first two of these changes result in a reduction of the error (before compen-

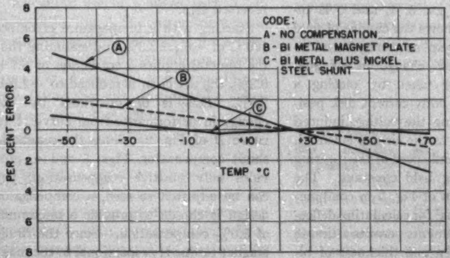
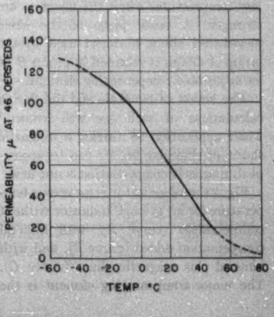


Fig. 2 (left). Tachometer indicator temperature error with and without compensation

Fig. 3 (right). Temperature-permeability characteristics of a magnetic shunt material



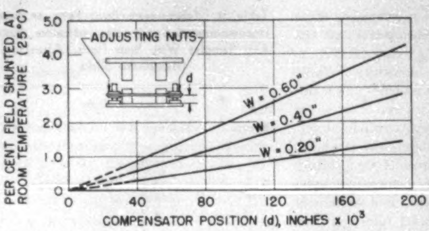


Fig. 4. Shunting action of magnetic shunts of different widths at 25 C

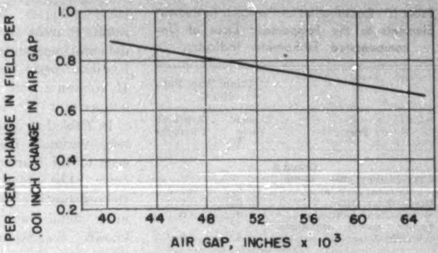


Fig. 5. Change of dragmagnet field strength with change in air gap

sation) from $-8.2\pm3.3\%$ to $-6.2\pm2.3\%$. The reason for the reduced error is the use of invar posts in group *B* as compared to the brass posts in group *A*.

The tachometers of group *A* were entirely compensated by the nickel-steel shunt[2] whose permeability-versus-temperature characteristics are approximately linear from −30 to +50 degrees centigrade (C); see Fig. 3. Because of the nonlinearity of this material over the specified temperature range (−55 to +70 C), it was never possible to reduce the temperature error of this tachometer design to the desired level. Attempts have been made to improve this method of compensation by using a shunt material which is linear over the desired temperature range; however, improved linearity is always accompanied by a reduction in slope, Fig. 3, which means that more field must be shunted to accomplish the same degree of compensation. This results in less flux in the air gap and a consequent reduction in torque gradient of the indicator. Because of these deficiencies, the method of compensation described in group *B* and illustrated in Fig. 1 was devised.

The improved compensation of group *B* uses a nickel-invar bimetal magnet plate, Fig. 1, which acts to close the air gap at elevated temperatures when the magnets are weak and to open the gap at low temperatures when the magnets are stronger. A small piece of the same nickel-steel shunt material used in the group *A* design is retained in group *B* as an adjustable compensating element.

The action of bimetals and the design calculations of such are well known.[1] Their application to moving a magnetic shunt or changing an air gap to accomplish magnetic compensation is not new.

Fig. 2 shows indicator error versus temperature for an average indicator without compensation (curve *A*), with bimetal compensation added (curve *B*), and with bimetal plus magnetic shunt (curve *C*). The major compensating element is the outer magnet plate made of a nickel-invar bimetal; see Fig. 1(B). Cast Alnico V magnets are pressed and soldered into the bimetal plate so that their position and the magnet air gap changes with temperature. Nickel-invar is used because both nickel and invar have relatively low magnetic reluctance.

The nickel-steel magnetic shunt, Fig. 1(A), is an adjustable compensator. It has been retained in the design to account for variables such as the eddy-current disk temperature coefficient (this parameter will vary with each new lot of material), the magnet temperature coefficient, the nickel-steel shunt coefficient, the variation in air gap adjustment, and the variation in bimetal coefficient. The position of the nickel-steel shunt is adjustable in the left-half of the dragmagnet, Figs. 1 and 4, to shunt a desired amount of field. Fig. 3 presents the average permeability-versus-temperature characteristic of the nickel-steel shunt material measured at 46 oersteds' field strength, a nominal value. The characteristic shown in Fig. 3 will hold over a wide range of shunt positions even though the field strength seen by the shunt is obviously a function of the shunt positions as indicated in Fig. 4.

The afore-mentioned compensation can be calculated and described mathematically. To demonstrate this, the design information of Figs. 3, 4, and 5 is included. Fig. 5 shows the change of field in the magnet air gap, as seen by the eddy-current disk, with change in air gap. These data were taken by placing a search coil in the eddy-current disk position and measuring the voltage induced in this search coil as the air gap is varied, the temperature and speed of the rotating dragmagnet being held constant. The design information of Fig. 5 in combination with equations[1] for calculating deflection of bimetal elements permits estimate of thermal activity and thickness of bimetal required to give a desired compensation. Since this procedure is well known, no detailed calculations are presented. However, allowance must be made for relative stiffness of bimetal and magnet-assembly posts in calculating change of air gap with temperature change. In the construction shown in Fig. 1, the bimetal plate is solidly clamped at three points to invar posts. The posts are so designed that they will bend with change of bimetal curvature with temperature change. For this reason their cross section is considerably reduced. If the posts were not flexible, they would seriously limit the bimetal action.

The data of Fig. 5 can also be used to estimate the change of dragmagnet field with the change in air gap caused by the relative expansion of dragmagnet assembly posts and magnets. Group *A* of Table I shows an estimated −1.0% change in torque (resulting from a −0.5% change in field) resulting from the relative expansion of brass posts and Alnico V magnets with a 100 C temperature change. Since the temperature coefficient of thermal expansion per C for Alnico V is about 11×10^{-6} compared to 20×10^{-6} for brass, and approximately zero for invar post material, the magnitude of the temperature error due to a change in air gap is about the same amount in group *B* as in group *A*, with reversed sign.

Starting with a temperature error per 100 C of $-6.2\pm2.3\%$ and assuming that the bimetal applies a correction of $3.7\pm0.2\%$, the net error is reduced to $-2.5\pm2.5\%$. This is demonstrated by the change in performance from curve *A* to curve *B* of Fig. 2. Since the magnetic shunt compensator, Figs. 3 and 4, provides only positive compensation, yet can be adjusted to zero, a compensator design is chosen to provide a maximum of 5.0% compensation. Since the dragmagnet torque is proportional to the magnetic field squared, the maximum field shunted should be approximately 2.5%. This will be accomplished by a shunt of 0.350-inch width, as indicated by the design data of Fig. 4.

It is important to note that the nickel-steel shunt in the improved design (group *B*, Table I) has an error of only 2.5±2.5% to compensate whereas in the old design (group *A*, Table I), the nickel-steel had to compensate the entire error of −8.2±3.3%. Having chosen a compensator of approximately the size required to compensate for the maximum error condition, it is possible with the design data of Figs. 3 and 4 to calculate the desired position of this compensator and the temperature error of the instrument over the complete temperature range.

Appendix I derives the following equation for % compensation in terms of the % field shunted at a given temperature due to the size and position of a shunt, Fig. 4, and the permeability-temperature characteristic of the nickel-steel shunt material, Fig. 3

$$\%\text{compensation} = 100 \times \left\{ \frac{\left[100-\phi_1\frac{\mu_2}{\mu_1}\left(\frac{\mu_1+K_AK_L}{\mu_2+K_AK_L}\right)\right]^2-(100-\phi_1)^2}{(100-\phi_1)^2} \right\} \tag{1}$$

For the shunt material used, Fig. 3, if ϕ_1 is defined as ϕ_{25} (the % field shunted at 25 C as described in Fig. 4) $\mu_1=\mu_{25}=50$ from Fig. 3, $K_A=5$, and $K_L=50$, with these parameters properly substituted equation 1 reduces to

$$\%\text{compensation}\Big|_{25}^{x} = 100 \times \left\{ \frac{\left[100-\phi_{25}\frac{u_x}{50}\left(\frac{300}{\mu_x+250}\right)\right]^2-(100-\phi_{25})^2}{(100-\phi_{25})^2} \right\} \tag{2}$$

which is an equation for calculating % compensation accomplished by this particular magnetic shunt in going from 25 C to some other temperature *x*. Checking, if $x=25$ C, then $\mu_x=50$ and equation 2 shows zero compensation, which it should.

Assuming use of a shunt with 0.280-inch width, Fig. 4, % compensation is calculated from 25 C to −30 C for a compensator position $d=90$ mils as follows: from Fig. 4, $\phi_{25}=1.0$ and, from Fig. 3, $\mu_{30}=118$ so that substituting these values into equation 2 results in

$$\%\text{compensation} = -1.87$$

This figure added to a known error at −30 C before compensation will give the temperature error of the compensated instrument with compensation. Similarly, compensation can be calculated for as many temperatures as desired so that points for a complete % error-versus-temperature curve can be determined.

Similarly, equation 2 with the data of Figs. 3 and 4 can be used to calculate shunt position to accomplish a desired % compensation. Referring to Fig. 2, curve *B*, assume that the % error is to be reduced to zero at −30 C. This requires a compensation of 1.6%. Substituting this value and $\mu_x=118$, and solving for ϕ_{25} by trial and error gives $\phi_{25}=0.86$.

Referring to Fig. 4, the compensator position corresponding to the % field shunted of 0.86 for the 0.280-inch-width shunt is $d=0.078$ inch.

Nomenclature

A_g = average area of the air gap
A_t = average area of the iron circuit
K_A = ratio of A_g to A_t
K_d = a constant
K_L = ratio of L_t to L_g
K_m = a constant MA_t/L_g
L_t = average length of the air gap between magnet and shunt
L_g = average length of the nickel-steel shunt circuit
M = magnetomotive force as seen by the shunt element
R_g = reluctance of the air gap between the magnet and the shunting element
R_t = reluctance of the nickel-steel shunt path
T = dragmagnet torque
T_1 = torque associated with ϕ_1
T_2 = torque associated with ϕ_2
ϕ = % field shunted
ϕ_1 = % field shunted at one temperature
ϕ_2 = % field shunted at another temperature
μ = permeability of the nickel-steel shunt
μ_1, μ_2 = shunt permeabilities at the two temperatures

Test Performance

As shown in Fig. 2, curve *A*, the total error of a typical uncompensated tachometer indicator is approximately −6.2% per 100 C. The addition of bimetal compensation reduces this error to approximately −2.5% per 100 C. Because of the nature of the bimetal, this type of compensation provides nearly linear correction over the entire temperature range, as shown by curve *B* of Fig. 2.

The remaining error is corrected by the nickel-steel shunt, Fig. 1(A), to provide the final temperature error shown by curve *C*. The sinusoidal shape of curve *C* is a result of the nonlinear characteristic of the nickel-steel shunt material, as shown by Fig. 3.

Fig. 6 compares the final temperature error of the old and the improved design. The nickel-steel shunt in the improved design has been adjusted to give a maximum error at −54 C which is approximately the same as the old design. This was done to provide the nearly flat compensation in the region from −10 C to +70 C, the range in which nearly all military flying is done. The compensation of the improved design may be adjusted to minimize temperature error over the full range of −55 C to +70 C. For this condition, maximum temperature error will be less than 0.5%.

A comparison of Table II with Table I shows the correlation between measured temperature error and estimated temperature error. The measured temperature error on five uncompensated indicators of the design represented by group *A* of Table I, brass posts and no bimetal, is an average of −9.5% per 100 C. The estimated error is −8.2±3.3% per 100 C; with the tolerance the maximum estimated error is −11.5%. Since all of the measured temperature errors are within the specified tolerance, the estimated results are justified.

Conclusion

The need for improved tachometer indicator performance made it desirable to analyze the source and magnitude of temperature errors in a dragmagnet eddy-current disk type of tachometer indicator. The results of this analysis are summarized in Table I.

Equations have been developed to calculate the % temperature compensation which can be accomplished with a nickel-steel shunt in the dragmagnet assembly. Design data necessary to the use of these equations are included in Figs. 3, 4, and 5.

A temperature-compensated dragmagnet design is described which is an improvement over an earlier design, as shown by the test performance of Figs. 2 and 6. The improved design is depicted in Fig. 1; it differs from the earlier design

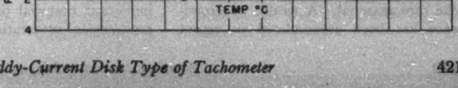


Fig. 6. Tachometer indicator temperature error with old style compensation and with improved compensation

in the use of a bimetal magnet plate, invar posts, and a reduced amount of nickel-steel shunt material.

Appendix. Derivation of Equation for Calculating % Compensation Accomplished by a Magnetic Shunt Element of the Type Described in Figs. 3 and 4

The reluctance of the magnetic circuit which shunts field away from the drag-magnet air gap is

$$R=R_i+R_g \tag{3}$$

From the physical dimensions of the parts, the following estimates are made

$$L_i=K_LL_g \tag{4}$$

$$A_g=K_AA_i \tag{5}$$

By definition

$$R_g=L_g/K_AA_i \tag{6}$$

$$R_i=\frac{K_LL_g}{A_i\mu} \tag{7}$$

From equation 3

$$R=R_i+R_g=L_g\Big/\frac{A_i(\mu+K_LK_A)}{(\mu K_A)} \tag{8}$$

By definition

$$\phi=M/R \tag{9}$$

Substituting equation 8 into equation 9

$$\phi=MA_i/L_g\frac{(K_A\mu)}{(K_AK_L+\mu)}=K_m\frac{(K_A\mu)}{(K_AK_L+\mu)} \tag{10}$$

$$\phi_1=K\frac{(K_A\mu_1)}{(K_AK_L+\mu_1)} \tag{11}$$

and

$$\phi_2=K\frac{(K_A\mu_2)}{(K_AK_L+\mu_2)} \tag{12}$$

T may be written in terms of % field shunted

$$T=K_d(100-\phi)^2 \tag{13}$$

By definition

$$\%\text{ compensation}=100\frac{\Delta T}{T}=100\frac{(T_2-T_1)}{(T_1)} \tag{14}$$

Substituting equations 11 and 12 into equation 13 and the resulting equation into equation 14

$$\%\text{ compensation}=100\times\left\{\frac{\left[100-\phi_1\frac{\mu_2}{\mu_1}\left(\frac{\mu_1+K_AK_L}{\mu_2+K_AK_L}\right)\right]^2-(100-\phi_1)^2}{(100-\phi_1)^2}\right\} \tag{15}$$

References

1. **Truflex.** *Thermostat Metals Catalog*, General Plate Company, Attleboro, Mass.
2. **Temperature Errors in Induction Watt-hour Meters**, I. F. Kinnard, H. T. Faus. *AIEE Transactions*, vol. 44, 1925, pp. 275–85.
3. **Recent Advances in Aircraft Tachometer Design**, R. G. Ballard, C. P. Hall. *AIEE Transactions (Electrical Engineering)*, vol. 63, Sept. 1944, pp. 646–48.
4. **The Magnetic-Drag Tachometer**, R. G. Ballard. *AIEE Transactions*, vol. 61, 1942, pp. 366–8.
5. Military Specification *MIL-I-7069*, Washington, D. C., Dec. 29, 1950.
6. **Aircraft Tachometer Indicator—An Analysis of Design Factors Affecting Starting Performance**, L. T. Akeley. *AIEE Transactions*, vol. 73, pt. II, 1954 (Jan. 1955 section), pp. 398–403.

Relation of D-C Magnetic Properties of Oriented 48-Per-Cent Nickel-Iron to Magnetic-Amplifier Performance

C. E. WARD
NONMEMBER AIEE

M. F. LITTMANN
ASSOCIATE MEMBER AIEE

TO IMPROVE the characteristics of cores for magnetic amplifiers, it is necessary for the metallurgist to isolate and measure those properties under his control as a material producer. The performance of a completed core depends on many factors, not all of which reflect actual magnetic properties of the core material itself. Many methods[1–6] have been used successfully to measure core performance under production conditions for matching and grading; similarly, several excellent papers[3,4,6,7] have been presented relating core performance to magnetic-amplifier design. While these dynamic methods are most valuable to the core producer and designer, their very advantage in including all the factors contributing to core performance has proved to be a handicap to one seeking to control factors which can be attributed only to the core material. A wealth of experience has been accumulated in recent years relating metallurgical factors such as chemical composition and processing to d-c magnetic properties. D-c magnetic tests have been widely used to investigate material variations because of their accuracy, reproducibility and the fact that they are unaffected by eddy currents. For example, to study the influence of an objectionable impurity it is important to control or, better, to eliminate other effects such as thickness and volume resistivity. The d-c test accomplishes this because it will reflect the changes caused by the chemical impurity without being influenced by other factors.

The object of this paper is to show what relationships exist between static and dynamic magnetic properties of joint significance to material producer and designer. A planned group of cores was prepared with controlled levels of such important material properties as thickness, crystal orientation, and coercivity. These properties were measured by d-c tests which were treated as the independent variables in this study. Using graphical correlation methods, the effects of these variables on dynamic properties were studied using standardized cores and carefully calibrated test circuitry.

Test Methods

A number of investigators have shown that core characteristics as determined by a self-saturating half-wave magnetic-amplifier tests are related to performance of various types of magnetic amplifiers. For this reason and because of its simplicity, the half-wave circuit was used to obtain the flux-control characteristics of the cores used in this study. In addition, a third method which has gained some prominence as a method of qualitatively evaluating magnetic-amplifier core performance was used. This method involves the normal dynamic flux-current loop.

In general, these methods and the corresponding parameters chosen conform to the standards recommended by the AIEE

Paper 55-455, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE-APS-AIME-CIT Conference on Magnetism and Magnetic Materials, Pittsburgh, Pa., June 14–16, 1955. Manuscript submitted March 14, 1955; made available for printing April 20, 1955.

C. E. Ward and M. F. Littmann are with the Armco Steel Corporation, Middletown, Ohio.

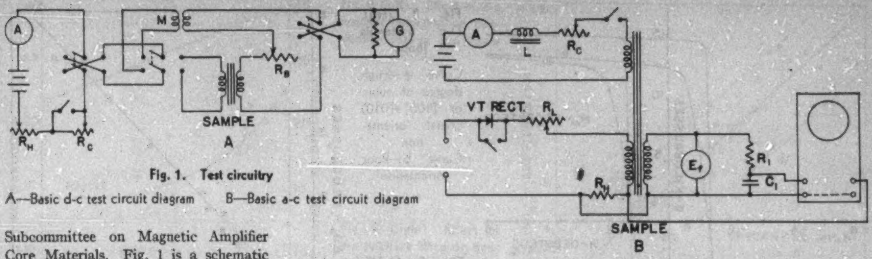


Fig. 1. Test circuitry

A—Basic d-c test circuit diagram B—Basic a-c test circuit diagram

Subcommittee on Magnetic Amplifier Core Materials. Fig. 1 is a schematic diagram of the basic test circuits used in determining the static and dynamic core characteristics.

The d-c ballistic test method, Fig. 1(A), is that given under ASTM Specification No. *A341-49*.[8] This type of test combines a high degree of accuracy and reproducibility with minimum effects of circuitry. It is considered the most satisfactory method for discerning small differences in basic magnetic properties.

The a-c test circuit shown in Fig. 1(B) combines equipment for obtaining both normal and magnetic-amplifier flux-current loops as well as the flux-control characteristic under constant-current reset conditions. Details of the method have been described by Dieterly.[9]

Test Parameters

Fig. 2(A) illustrates a typical d-c hysteresis loop with the test points of interest noted. Peak magnetizing forces of 1 and 10 oersteds [about 0.8 and 8 ampere-turns per centimeter (amp-turns/cm)] were used for obtaining the loops.

Normal flux-current loops as shown in Fig. 2(B) were obtained with sinusoidal flux conditions at peak excitations of 0.8 and 8 amp-turns/cm. Of prime interest are the a-c "coercive" force values obtained from these loops.

Conditions for the magnetic-amplifier tests were as follows. At zero control current or bias the load resistance was adjusted for a load current swing corresponding to 15 amp-turns/cm excitation in the core. At cutoff the input voltage was adjusted to give a flux swing in the core like that for a major loop with a peak excitation of approximately $2(NI)_c$. Fig. 2(C) illustrates the two end conditions under which the cores were operated as magnetic amplifiers. Fig. 2(D) illustrates a typical flux-control characteristic with the major points of interest detailed.

Preparation of Cores

The material used was a 48-per-cent nickel-iron alloy melted commercially in 10-ton arc furnaces. Strip was rolled to standard final thicknesses varying from 6 mils to 1/4 mil. Cores were prepared from cold-rolled tape slit to 1/4-inch width and wound into toroidal cores $1^5/_8$ inches inside diameter by 2 inches outside diameter using magnesia insulation between wraps. The fabricated cores were annealed in dry hydrogen and then placed into molded boxes to protect them from stresses during the application of copper windings. To minimize effects other than material properties all the cores were wound with a primary of 600 turns, a secondary of 300 turns, and a control winding of 50 turns. All the properties were measured at room temperature.

Selection of Material Properties Studied

From an empirical point of view the basic variations possible in soft ferromagnetic materials (free from mechanical stresses) may be summarized as follows:

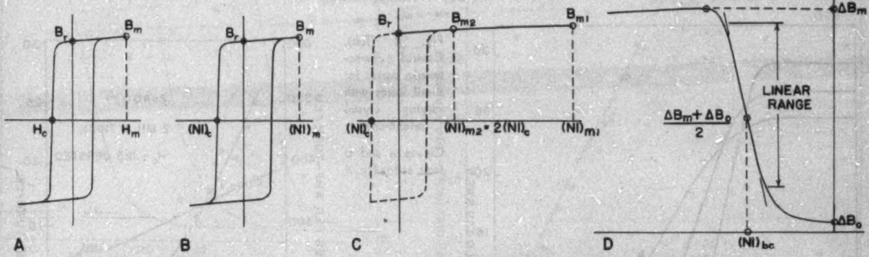


Fig. 2. Nomenclature

A—D-c hysteresis loop

B_m = maximum induction, kilogausses

B_r = residual induction, kilogausses

H_c = coercive force, oersteds

H_m = magnetizing force, oersteds

B—A-c flux-current loop

$(NI)_m$ = maximum a-c excitation, amp-turns/cm

$(NI)_c$ = a-c coercive force, amp-turns/cm

C, D—Magnetic-amplifier flux-control curve

ΔB_0 = flux-density change at zero bias, kilogausses

ΔB_m = maximum flux-density change at cutoff, kilogausses

G = gain, defined as ratio of linear flux change to corresponding change in bias. Linear range is that ΔB over which the d-c bias does not vary over 0.004 amp-turn/cm from linearity

$(NI)_b$ = bias or d-c control magnetizing force, amp-turns/cm

$(NI)_{bc}$ = bias requirement corresponding to d-c coercive force, amp-turns/cm. (NI)bc is measured where $\Delta B=(\Delta B_0+\Delta B_m)/2$

$(NI)_{m1}$ = maximum a-c excitation at zero bias, amp-turns/cm

$(NI)_{m2}$ = maximum a-c excitation at cutoff, amp-turns/cm

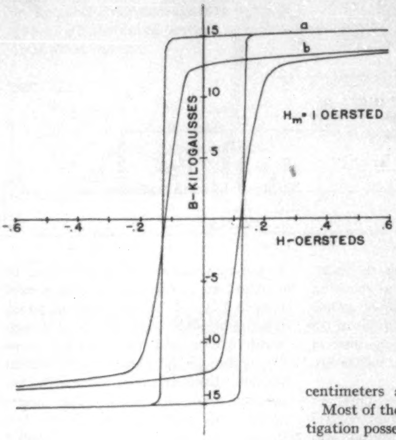


Fig. 3 (left). D-c hysteresis loops

Curve a—High degree of cubic or [100] (010) crystal orientation
Curve b—Poor orientation

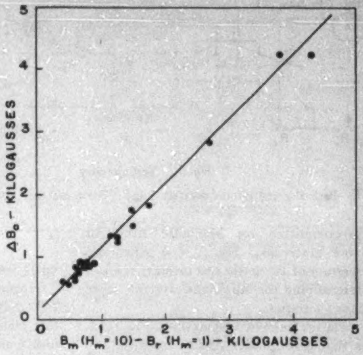


Fig. 5 (right). Effect of crystal orientation on flux - density change at zero bias

1. Gross chemical composition.
2. Degree of freedom from chemical impurities.
3. Crystal or grain orientation.
4. Tape thickness.
5. Annealing effects.

Gross chemical composition and hence saturation and volume resistivity were fixed by choice of the 48-per-cent nickel-iron alloy. A typical saturation value of ferric induction (**B** minus **H**) is approximated by an induction of about 15,250 gausses reached at a peak magnetizing force of 10 oersteds. Volume resistivity was found to be some 44 to 47 microhm-centimeters at 25 degrees centigrade.

Most of the cores studied in this investigation possessed a high degree of crystal orientation as found in 48 Orthonik. Such material is characterized by a rectangular hysteresis loop with a residual induction $\mathbf{B}_r$ in excess of 13.50 kilogausses for a tip magnetizing force $\mathbf{H}_m$ of 1 oersted. Compared to relatively "unoriented" material, the oriented material tends to have higher coercive force for thicknesses of 2 to 6 mils although the reverse tends to be true for thicknesses below 1 mil. The coercive force attainable in 48 Orthonik usually decreases with thickness down to about 1 mil. For this reason a set of cores was specially prepared from thicknesses of 1/4 mil to 6 mils in order to study the effect of thickness for a constant d-c coercive force. Similarly the effect of coercive force was studied for varying degrees of orientation keeping all other factors as constant as possible. This was achieved by variations in melting, rolling, and final annealing treatments.

Melting, processing, and annealing all work together to produce the magnetic properties of the core material for a given tape thickness. Together with dimensional effects, insulation between wraps, externally applied stresses, and test circuitry (especially for dynamic tests), they ultimately account for the properties observed in the cores. The primary interest in this study is the influence of those properties characteristic of the core material itself.

Effect of Crystal Orientation

Even though 48-per-cent nickel-iron is not strongly anisotropic, the effects of

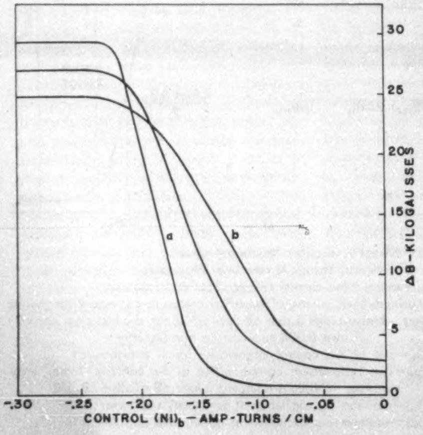


Fig. 4 (left). Control characteristic curves for 2-mil cores with varying crystal orientation

Curves a and b same as for Fig. 3

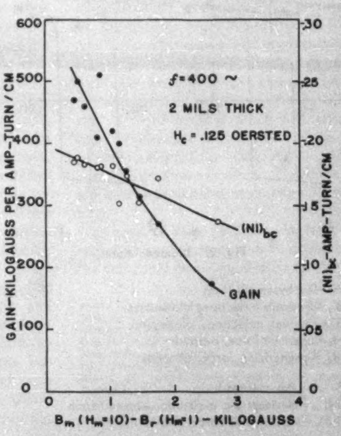


Fig. 6 (right). Effect of crystal orientation on gain and control bias coercive force

Vol
Pt.
195

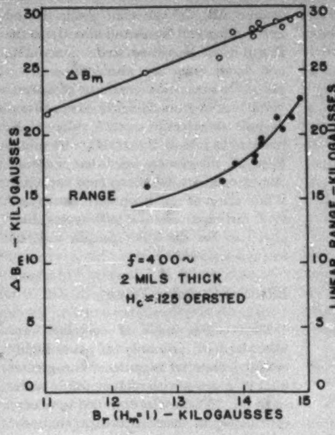
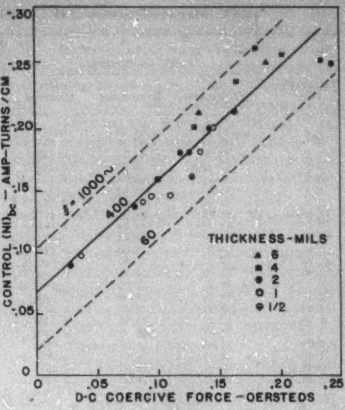


Fig. 7 (left). Effect of d-c residual induction on flux-density change and linear range

Fig. 9 (right). Effect of d-c coercive force on control bias coercive force for 48 Orthonik

NOTE: The data have been omitted on the 60- and 1,000-cycle curves for the sake of clarity. Complete data are listed in Table I



crystal orientation nevertheless are rather striking, both on d-c magnetic properties as well as on dynamic properties evidenced by control characteristic curves. The type of hysteresis loop made possible by a high degree of perfection of [100] (010) crystal orientation is the familiar rectangular loop shown in Fig. 3, curve *a*. By contrast, a poorly oriented material with nearly the same coercive force is shown in Fig. 3, curve *b*. A convenient but sensitive index of the degree of crystal orientation is the residual induction $\mathbf{B}_r$ for a tip magnetizing force which substantially saturates the material. A highly oriented material such as 48 Orthonik will have a $\mathbf{B}_r/\mathbf{B}_m$ ratio of about 0.96 to 0.98 at $\mathbf{H}_m=1$ oersted. The $\mathbf{B}_r$ value changes only very little when $\mathbf{H}_m$ is increased to 10 oersteds, although the $\mathbf{B}_m$ value will increase by an additional 1 or 2 per cent. On the other hand, changes in orientation can cause the $\mathbf{B}_r$ value to drop as low as 10 kilogausses. For such a material the value of $\mathbf{B}_m$ at $\mathbf{H}_m=1$ oersted may be only about 13 kilogausses and increases to 15 kilogausses at $\mathbf{H}_m=10$ oersteds. Use of the difference between $\mathbf{B}_m$ at $\mathbf{H}_m=10$ and $\mathbf{B}_r$ at $\mathbf{H}_m=1$ rather than $\mathbf{B}_r$ alone as an index of orientation eliminates the error resulting from small

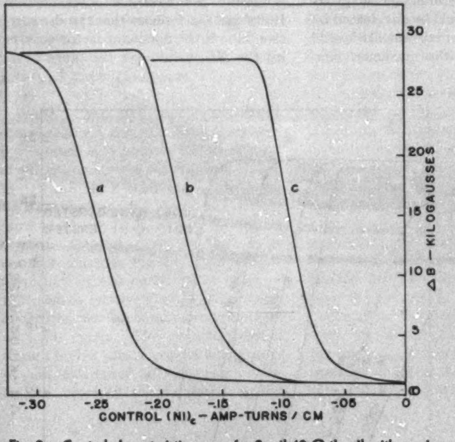


Fig. 8. Control characteristic curves for 2-mil 48 Orthonik with varying d-c coercive force

D-C Properties at $H_m=1$ Oersted

	B_m, Gausses	B_r	H_c, Oersteds
Curve a	14,900	14,500	0.240
Curve b	15,000	14,600	0.124
Curve c	14,900	14,050	0.031

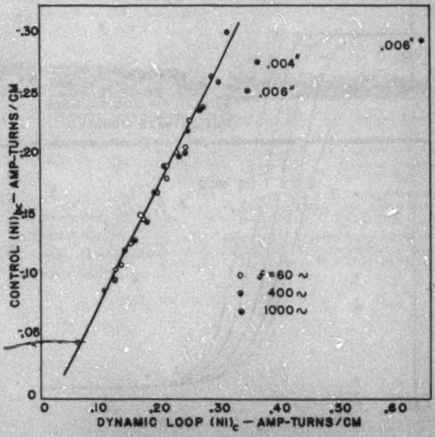


Fig. 10. Relation between dynamic loop coercive force and control bias coercive force for 48 Orthonik

Table I. Magnetic Properties of 48 Orthonik Cores (Refer to Fig. 9)

Thickness, Mils	$H_m = 1$ Oersted			$H_m = 10$ Oersteds	$(NI)_{bc}$, Amp-Turns/Cm		
	B_m, Kilogausses	B_r, Kilogausses	H_c, Oersteds	B_m, Kilogausses	60 Cycles	400 Cycles	1,000 Cycles
6.0	14.75	14.45	0.191	15.35	0.203	0.255	0.207
6.0	15.05	14.55	0.134	15.35		0.213	
3.8	15.00	14.75	0.203	15.40	0.208	0.262	
3.8	15.15	14.90	0.165	15.40	0.182	0.240	0.279
3.9	15.00	14.65	0.131	15.45	0.148	0.202	
1.9	14.85	14.45	0.245	15.40	0.230	0.253	0.272
1.8	14.90	14.55	0.236	15.40	0.221	0.256	0.300
1.9	15.10	14.85	0.177	15.50	0.194	0.267	0.304
2.3	14.75	14.30	0.164	15.15	0.172	0.215	
2.2	15.15	14.90	0.143	15.50		0.202	
1.9	15.00	14.60	0.124	15.40	0.128	0.182	0.222
2.0	15.05	14.80	0.118	15.30		0.183	
1.9	14.90	14.30	0.100	15.35	0.108	0.160	
2.1	15.10	14.70	0.081	15.40	0.087	0.137	0.174
2.1	14.90	13.90	0.028	15.40	0.047	0.090	0.124
1.05	15.05	14.85	0.147	15.35	0.155	0.202	0.238
1.00	15.15	14.85	0.135	15.45	0.136	0.182	0.220
1.10	14.85	14.40	0.110	15.20	0.112	0.147	0.172
1.05	14.95	14.60	0.096	15.25	0.106	0.147	
1.00	15.15	14.75	0.088	15.50	0.102	0.142	0.172
0.95	15.05	14.50	0.037	15.45	0.044	0.098	0.132
0.46	15.20	14.90	0.129	15.55		0.162	0.188

differences in magnetic saturation between batches of material.

Fig. 4 illustrates the control characteristic curves at 400 cycles for 2-mil cores with nearly identical coercive force but widely varying degrees of crystal orientation. Control curves *a* and *b* of Fig. 4 were obtained on the same cores whose d-c loops are shown in Fig. 3, curves *a* and *b* respectively. The effect of crystal orientation on the dynamic characteristics is very important since $\Delta \mathbf{B}_0$, gain, NI_{bc}, linear range of gain, and $\Delta \mathbf{B}_m$ are all affected, as will be shown. A quantitative picture of the effect of orientation on $\Delta \mathbf{B}_0$ is shown in Fig. 5. This relation is independent of tape thicknesses and test frequencies studied. Fig. 6 shows the strong effect of orientation on gain. The effect of orientation on the bias coercive force NI_{bc} as shown in Figs. 4 and 6 is an interesting confirmation of the observation that rapid changes in magnetization in rectangular-loop materials are associated with a time lag or higher energy requirements than would be anticipated by conventional eddy-current calculations.[4,9] Fig. 7 shows that the maximum flux-change $\Delta \mathbf{B}_m$ of the control curve and linear range can be related directly to the $\mathbf{B}_r$ value for $\mathbf{H}_m = 1$ oersted. Since $\Delta \mathbf{B}_m$ and linear range are also influenced by magnetic saturation (because of normal variations in gross chemical composition) as well as orientation, the value of $\mathbf{B}_r$ itself rather than $\mathbf{B}_m$ at $\mathbf{H}_m = 10$ minus $\mathbf{B}_r$ at $\mathbf{H}_m = 1$ gives a somewhat more accurate estimate of these two variables. While these effects have been illustrated for 2-mil tape tested at 400 cycles, they hold true for the other gauges and test frequencies studied.

Effect of Coercive Force

The studies made of coercive force were limited primarily to the highly oriented material because of the greater interest in cores with rectangular hysteresis loops. The results obtained were very conclusive in demonstrating that d-c coercive force $\mathbf{H}_c$ is a very important factor governing NI_{bc} even over a wide range of thicknesses and frequencies. For the sake of comparison, the effects of $\mathbf{H}_c$ for $\mathbf{H}_m = 1$ oersted are shown at 400 cycles for 2-mil cores in Fig. 8. In addition to controlling the value of NI_{bc}, d-c coercive force also affects gain. Although this effect on gain is quite prominent at 60 cycles, it diminishes at 400 cycles and is hardly noticeable at 1,000 cycles. However, Fig. 9 shows that the d-c coercive force is the dominant factor governing the NI_{bc} value over the range of 60

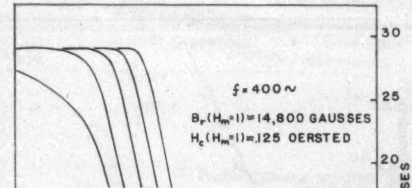


Fig. 11. Control characteristic curves for 48 Orthonik 1/2 to 6 mils thick having similar d-c properties

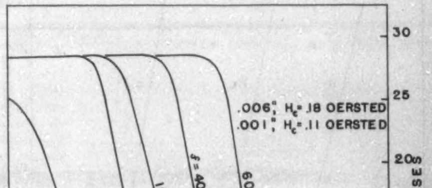


Fig. 12. Variations of control characteristic curves with frequency for typical 6-mil and 1-mil 48 Orthonik

to 1,000 cycles. The data on which Fig. 9 is based are shown in Table I. The data were used without any corrections for other factors except that the group was limited to cores with values of $\mathbf{B}_m$ at $\mathbf{H}_m=10$ minus $\mathbf{B}_r$ at $\mathbf{H}_m=1$ not over 1.00 kilogausses. Within the group, cores with lesser thickness, lower magnetic saturation, or less pronounced crystal orientation tend to have lower NI_{bc} values for a given d-c coercive force.

It is interesting to note that the relation is linear and that the intercept of the curve could be described as a sort of "threshold" value of NI_{bc}, which is peculiar to this alloy but also depends on frequency and to a smaller degree on thickness. Reference to Fig. 4 indicates that this threshold value of NI_{bc} also depends to a considerable degree on crystal orientation. This confirms the results reported by Huhta[9] who found that a higher magnetizing force is required for the same rate of flux change in materials with the more rectangular hysteresis loop.

Because many are familiar with a-c coercive force NI_c obtained from dynamic flux-current loops, the relation between NI_c and NI_{bc} is shown in Fig. 10 for 48 Orthonik. The relation is very close and independent of thickness and frequency apparently until eddy-current effects become dominant as for the 6-mil specimens at frequencies of 400 and 1,000 cycles.

Effect of Tape Thickness

While d-c properties usually attainable commercially vary with tape thickness, it was possible to prepare and study a set of 48 Orthonik cores varying from 6 mils to 1/2 mil thick with nearly the same $\mathbf{B}_r/\mathbf{B}_m$ ratio of about 0.98 and d-c coercive force of about 0.125 oersted for $\mathbf{H}_m=1$ oersted. Control curves for these cores are reproduced in Fig. 11, again using a frequency of 400 cycles. The effect of thickness is relatively small under these conditions, but becomes more prominent at 1,000 cycles. The general trend is toward higher gain, a greater linear range of gain, and lower NI_{bc} as thickness decreases, provided the d-c coercive force remains the same. A sacrifice in performance is experienced if a core is used at frequencies above that appropriate to its tape thickness. This is illustrated in Fig. 12 where typical 6-mil core shows a severe drop in gain and increased bias requirements as $\Delta\mathbf{B}_m$ is approached at 400 and especially 1,000 cycles. This is because there is not enough time for the flux to be completely reset during the half-cycle when the bias current is operative. By contrast, the control characteristic curves for a typical 1-mil core are also shown in Fig. 12. With the advantage of lower d-c coercive force and the ability to be reset more quickly, the 1-mil material still performs well even at 1,000 cycles.

Summary

The basic magnetic properties of 48-per-cent nickel-iron measured by static tests have been found to have a strong effect on dynamic properties pertinent to magnetic-amplifier operation. The principal factors affecting magnetic-amplifier core performance are crystal orientation and coercive force.

For toroidal cores of a given thickness, there is an improvement in most characteristics desirable for magnetic amplifiers with the higher degree of crystal orientation. Principally these involve decreased flux change at zero control current, better gain and linearity of gain, and a higher, total flux swing in the core. However, there is a tendency for increased bias current requirement to initiate a flux change in the core.

The second important factor is d-c coercive force which governs the value of bias control current required. This is a linear relation above a certain threshold value which, for the alloy studied, varies with frequency and to some extent with orientation and thickness. Low coercive force is associated with high gain, especially at 60 cycles; at 400 cycles the effect is substantially reduced.

The effects of orientation and coercive force are very similar for thicknesses in the range of 1/2 to 6 mils and for the frequencies of 60 to 1,000 cycles studied. When thicknesses are compared for similar coercive force and orientation, the thinner material shows better gain and increased linear range of gain even under conditions where eddy-current effects, as normally interpreted, are unimportant. The effect of thickness is quite pronounced when the flux does not completely reset during the half-cycle periods when the d-c bias is applied. For a thickness of 4 mils this becomes quite evident at 1,000 cycles and for 6 mils at 400 cycles.

Dynamic tests such as the constant-current reset method and flux-current loop method provide valuable information for core matching, core grading, and the design of magnetic amplifiers. On the other hand, d-c magnetic tests most accurately reflect magnetic characteristics familiar to the metallurgist. It is hoped that the quantitative data relating basic material characteristics with performance measured by dynamic tests will enable material producer, core fabricator, and designer to co-operate more effectively in improving core materials for magnetic amplifiers.

References

1. Dynamic Hysteresis Loop Measuring Equipment, **H. W. Lord.** *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 269–72.

2. Magnetic Amplifier Gapless Core Tests, **J. R. Conrath.** *Electronics*, New York, N. Y., Nov. 1952, pp. 119–21.

3. Evaluation of Core Materials for Magnetic Amplifier Applications, **R. D. Teasdale, H. R. Brownell.** *Convention Record of the IRE*, Institute of Radio Engineers, New York, N. Y., Mar. 23, 1954.

4. Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers, **R. W. Roberts.** *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 682–90.

5. Pulse Response Characteristics of Rectangular-Hysteresis Loop Ferromagnetic Materials, **Joseph Wylen.** *Ibid.*, vol. 72, pt. I, Nov. 1953, pp. 648–56.

6. Evaluation of Core Materials for Magnetic Amplifiers, **D. C. Dieterly.** *Electrical Manufacturing*, New York, N. Y., vol. 51, no. 1, Jan. 1953, p. 68, vol. 51 no. 2, Feb. 1953, p. 124.

7. Techniques for Measuring Cascaded Self-Saturating Magnetic-Amplifier Performance, **Henry Kaplan, Gerald Wolff.** *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 581–84.

8. ASTM Standards 1952. *ASTM Specification No. A341–49*, American Society for Testing Materials, Philadelphia, Pa., pt. I, 1952, p. 1309.

9. Flux Resetting Characteristics of Several Magnetic Materials, **Hoobert Huhta.** *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 111–14.

Voltage Gain of a Resonant Dielectric Amplifier

E. A. SACK
ASSOCIATE MEMBER AIEE

G. W. PENNEY
FELLOW AIEE

Synopsis: The gain of the resonant-circuit dielectric amplifier can be obtained directly from the slope of a characteristic which is called the tuning curve. When the carrier voltage across the nonlinear capacitor is small, the shape of the tuning curve can be computed by linear circuit analysis. This paper considers the shape of the tuning curve when the carrier voltage across the nonlinear capacitor is so large that nonlinear resonance is encountered.

THE dielectric voltage amplifier shown in Fig. 1 utilizes a nonlinear capacitor as part of a series resonant circuit. This circuit is excited by a power or carrier source v, whose amplitude and frequency are constant. The signal to be amplified is applied to the nonlinear capacitor C_N through a choke L_B which acts to prevent the flow of carrier current about the control loop. Capacitor C_L serves to prevent the flow of control current about the carrier loop.

The capacity of the nonlinear element is varied by the signal. A change in the capacity in the resonant loop alters the tuning and since the carrier frequency is constant the circulating carrier current is modulated by the signal. The carrier voltage appearing across inductance L is likewise modulated. When this voltage is demodulated by the rectifier and carrier filter, an output is obtained which, for a properly designed circuit, has similar wave shape but larger amplitude than the input signal. Thus a voltage gain is realized.

In an earlier paper the authors presented a method for the calculation of the low-signal frequency voltage gain of the amplifier.[1] That method utilizes the results of a linear circuit analysis and yields satisfactory results only so long as the circulating current in the resonant loop remains small. It is the purpose of this paper to present an analysis which relaxes this restriction in order to obtain a better approximation to the low-signal frequency voltage gain under typical operating conditions. In a paper on the frequency response of a resonant dielectric amplifier,[2] the authors discussed a solution which gives the voltage gain of the amplifier at any signal frequency in terms of the low frequency gain. Although this solution is primarily limited to the case where the circuit is in essentially linear operation, it can yield a very rough approximation to the signal frequency response of the amplifier when the circuit operation is highly nonlinear.

Tuning Curve

The low-signal frequency gain of the amplifier can be obtained directly from a characteristic which will be called the tuning curve. This characteristic relates the steady-state circulating current in the resonant carrier loop to the fixed or slowly varying control potential applied to the nonlinear capacitor. A typical tuning curve is shown in Fig. 2. As the signal varies about the fixed bias point, the variation in the amplitude of the circulating current can be traced on the curve. The output of the amplifier is proportional to this variation. Therefore, the low-signal frequency gain is proportional to the slope of the tuning curve in the region of operation. Knowledge of the shape of the curve is of fundamental importance both in adjustment of the amplifier and in computation of the gain.

At low levels of circulating current the tuning curve of the resonant carrier loop has approximately the shape predicted by linear analysis. As the current is increased, however, the curve becomes distorted because of the phenomenon of nonlinear resonance. To predict the shape of the tuning curve at higher circulating current levels it is first necessary to select a function to approximate the the electrical terminal characteristics of the nonlinear capacitor. Once this function has been chosen, the tuning curve can be obtained by solution of the nonlinear differential equation for the resonant loop.

Equivalent Nonlinear Capacitor

At present the material most commonly used as a dielectric for nonlinear capacitors is barium strontium titanate. This material is frequently called ferroelectric because of the similarity between the curves of magnetic induction versus magnetic intensity of ferromagnetic materials and the electric displacement versus gradient (D-E) curves of the titanate. Fig. 3 shows a major hysteresis loop for a barium-strontium-titanate capacitor.

The following is a list of factors other than control potential which appreciably influence the electrical behavior of a certain commercial supply of titanate dielectrics:

1. Temperature.
2. Test frequency.
3. Previous electrical history.
4. Electrode pressure.
5. Fluid of immersion.

In addition, the materials age, and there appears to be some random drifting of the characteristics. For the purposes of analysis no attempt will be made to use a multivalued charge versus potential function in order to represent the hysteresis accurately. Instead, the real nonlinear capacitor will be replaced by a lossless nonlinear capacitor in series with a linear resistor chosen to yield the same total energy loss per carrier cycle. This procedure is similar to that employed by other investigators, and the validity of the approximation relies on the fact that the energy lost per cycle is but a few per cent of the peak energy stored in the capacitor.[3-7]

Before selecting the single-valued curve

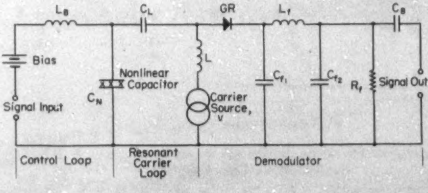


Fig. 1. Resonant-circuit dielectric amplifier

Paper 55-71, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 14, 1954; made available for printing March 1, 1955.

E. A. Sack is with the Westinghouse Electric Corporation, East Pittsburgh, Pa., and G. W. Penney is with the Carnegie Institute of Technology, Pittsburgh, Pa.

This work has been done in part under Office of Naval Research Contract *N9-onr-30308* and in part by E. A. Sack in partial fulfillment of the requirements for the Ph.D. degree at Carnegie Institute of Technology.

Vol.
Pt. 1
195

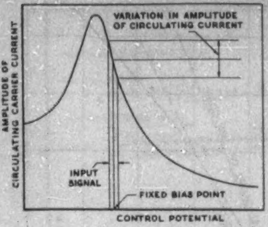


Fig. 2 (left). A tuning curve

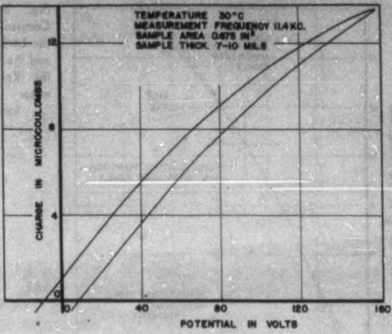


Fig. 3 (right). A dynamic major hysteresis loop for a barium-strontium-titanate capacitor

to approximate the dynamic charge versus voltage characteristics of the nonlinear capacitor it is necessary to review the physical situation in the carrier circuit. Since the nonlinearity of the titanate capacitor is small and since the resonant loop will tend to restrict the flow of off-resonance harmonics, both the circulating current and the voltages across circuit elements will be nearly sinusoidal. This has been verified by experiment. As previously stated, the object of this analysis is to determine the amplitude of the circulating current in the carrier loop of Fig. 1 as a function of the control potential for a given carrier source amplitude. For each value of control potential the charge and voltage on the nonlinear capacitor will describe a different steady-state dynamic minor loop. If the shape and location of the minor loop were known for a given set of circumstances there would be no question as to the relation between charge and voltage for purposes of analysis, but, of course, there would also be little need for analysis. It is therefore necessary to establish some basis for the choice of an average single-valued charge versus voltage curve for the equivalent lossless nonlinear capacitor.

Corresponding to every value of control potential on the nonlinear capacitor there will be a different family of minor loops. At first glance it would appear to be reasonable to select the locus of the tips of the minor loops to be the average dynamic charge versus voltage curve. However, the loci of the tips of the minor loops for different control potentials need not coincide. Therefore, it would seem to be indicated that a different function, or at least different numerical coefficients, must be used in the display of the nonlinearity for each value of control potential. Since the stated purpose of this analysis is the determination of the variation in circulating current as a function of the slowly varied control potential, it would hardly be convenient to use a different function or coefficients for each value of control potential. An alternative method of representation will therefore be selected.

Fig. 4 shows a curve drawn through the tips of the major hysteresis loops for a particular nonlinear capacitor. For convenience this curve will be called the locus of major loop tips (LMLT) curve. Fig. 5 compares the LMLT curve and the locus of the tips of a particular family of minor loops. Although the curves coincide for the small and medium size minor loops, the tips of the larger minor loops fall off of the LMLT curve, especially above the fixed control point. The shape of the curves is sufficiently similar, however, that for the purposes of this approximate analysis the LMLT curve will be selected as being a convenient (single-valued) approximation to the dynamic charge versus voltage characteristics of the capacitor. In the analysis to follow it will be as-

---

Nomenclature

A = area of the parallel plate nonlinear capacitor
a_3, b_3 = coefficients of the cubic approximation to D versus E
C_L = linear blocking capacitor in the carrier loop
C_N = nonlinear capacitor
D = electric displacement in the nonlinear dielectric material
E = electric gradient in the nonlinear dielectric material
G = slope of the β versus V_c curve in the region of operation
I = peak of the fundamental component of circulating carrier current
i_c = instantaneous current in the control loop
i_L = instantaneous current in the carrier loop
K_1, K_2 = constants in the hyperbolic approximations to q_N versus v_N
L = inductance in the carrier loop
l = thickness of the parallel plate nonlinear capacitor
$Q = (\omega_c L)/R$
Q_{N0} = coefficient of the tanh approximation to q_N versus v_N
q_c = charge flowing in the control loop
q_L = instantaneous charge on the linear blocking capacitor
q_N = charge on the nonlinear capacitor
q_{NT} = time-varying component of q_N
R = resistance in the carrier loop
t = time
V = peak of the carrier supply potential
V_c = control potential
V_{N0} = coefficient of the sinh approximation to q_N versus v_N
v = instantaneous carrier supply potential
v_N = potential on the nonlinear capacitor
VA = voltage amplification of the dielectric amplifier of Fig. 1
α = constant component of ρ_1
β = amplitude of the fundamental frequency component of ρ_1
ω_c = carrier angular frequency
ϕ_c = carrier phase angle
$(a_3), (b_3)$ = coefficients of the cubic approximation to q_N versus v_N
$(a_5), (b_5), (c_5)$ = coefficients of the quintic approximation to q_N versus v_N

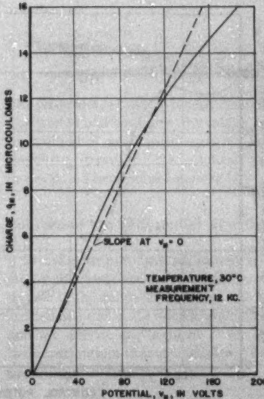


Fig. 4. LMLT curve

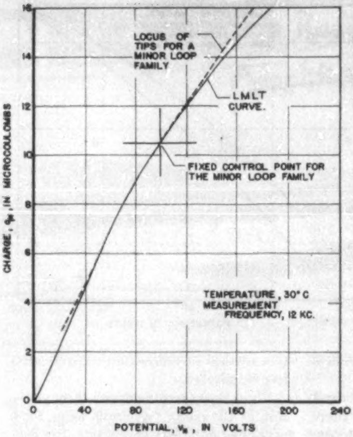


Fig. 5 (left). Comparison of the LMLT curve and the locus of the tips of a minor loop family

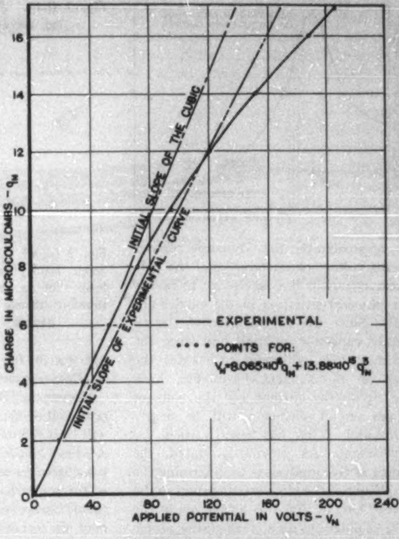


Fig. 7 (right). Cubic representation compared with the LMLT curve

sumed that the equivalent lossless nonlinear capacitor adheres to this charge versus potential curve for all values of control bias. A solution based upon this representation is necessarily a solution to a somewhat idealized problem.

There is reasonable doubt that greater accuracy of representation than that afforded by the LMLT curve is justified at present. As indicated earlier, the characteristics of the nonlinear capacitors which have been available are found to undergo considerable variation with time. Kazarnovski has reported observation of similar variations.[8]

The unpredictability in the characteristics due to variations with time is of the same order of magnitude as the error introduced by selection of the LMLT curve to represent the average charge versus voltage characteristic about any fixed control point. In Fig. 5 the tip of the minor loop for a peak swing of 100 volts deviates by 5.5 per cent from the LMLT curve. In a test with carefully controlled conditions, extending over a period of 6 hours, the tip of a loop for 164 volts' peak swing wandered over a range that is 5.4 per cent of the value during initial measurement. The inconvenience inherent in choosing a different function to represent the dynamic charge versus voltage curve for each control potential does not appear to be justified under these circumstances.

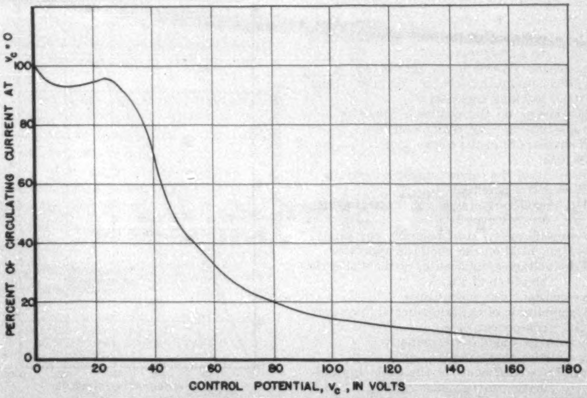


Fig. 6. Experimental tuning curve showing the effect of an initial polarization phenomenon

Analytical Display of Nonlinear Charge Versus Voltage Characteristic

Various investigators have used the cubic $v_N=(a_3)q_N+(b_3)q_N^3$ to approximate the single-valued curve which they picked for the nonlinear capacitor.[5-7] Others have used the hyperbolic functions $v_N = V_{N0} \sinh (K_1 q_N)$ and $q_N = Q_{N0} \tanh (K_2 v_N)$.[1,3,4] Either the cubic or the hyperbolic functions will represent the shape of the LMLT curve at higher values of control potential. However, neither of them can take into account the initial polarization effect near the origin. This polarization phenomenon is apparent in the LMLT curve of Fig. 4 where the slope initially increases before it begins to decrease. A similar effect can be observed on mean magnetization curves for ferromagnetic materials.

Even though the initial polarization effect is small, it gives rise to a very noticeable change in circulating current in the carrier loop of Fig. 1 as the control potential is increased from zero. Fig. 6 is an experimental tuning curve showing the steady-state circulating current versus

Vol
Pt.
195

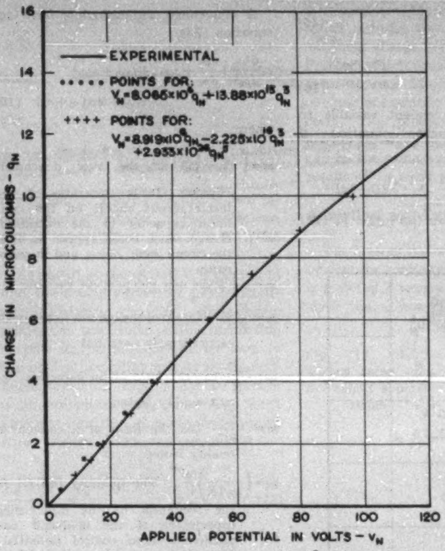


Fig. 8 (left). Quintic representation compared with the cubic and the LMLT curve

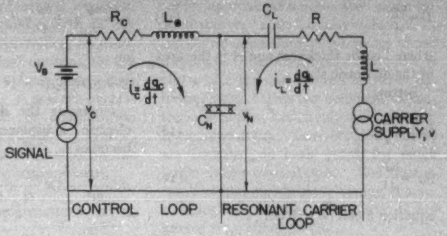


Fig. 9 (above). Carrier and control loops of the resonant-circuit dielectric amplifier

control potential. An analysis based on the cubic or hyperbolic approximations could never predict the initial dip in current.

So long as the instantaneous potential across the nonlinear capacitor does not range over that part of the LMLT curve near the origin, the cubic approximation can represent the average charge versus voltage characteristic quite adequately. This condition is often satisfied in the resonant-circuit dielectric amplifier, and results obtained from an analysis based on the cubic can be of value provided that the inherent restriction is held in view. Fig. 7 compares the LMLT curve for a test sample with the function

$$v_N=8.065\times10^6 q_N+13.88\times10^{15} q_N^3 \tag{1}$$

The quintic

$$v_N=8.919\times10^6 q_N-2.225\times10^{16} q_N^3+2.933\times10^{26} q_N^5 \tag{2}$$

can be made to approximate the LMLT curve in the region of initial polarization. This analytical representation can be used when the instantaneous potential on the nonlinear element varies over the curve in the vicinity of the origin. It has not been possible to choose the coefficients of equation 2 so that the function conforms to the LMLT curve at higher as well as lower values of applied potential. The LMLT curve, the quintic, and the cubic approximations are compared in Fig. 8.

Circuit and Differential Equations

For the purposes of analysis it is generally assumed that the demodulator has only a negligible effect on the resonant loop. An improvement on this assumption is that the loading effects of the demodulator and any output load can be approximated by an equivalent resistance added in series with the resonant loop. In either case the circuit to be analyzed reduces to that shown in Fig. 9. All of the elements of the circuit except C_N are assumed to be linear. Since the analysis is limited to the steady-state behavior of the resonant loop with a fixed control potential V_c, the only current flow about the signal loop will be of the carrier frequency and its harmonics. In the dielectric amplifier, however, L_B and R_c are chosen so that $i_c<<i_L$. Under these circumstances the equivalent load resistance for the nonlinear capacitor is completely assigned to the carrier loop and is included in R.

The differential equation for the carrier loop is

$$v=L\frac{d^2q_L}{dt^2}+R\frac{dq_L}{dt}+\frac{1}{C_L}q_L+v_N \tag{3}$$

where q_L is the charge on the linear capacitor C_L. For purposes of analysis it is convenient to assign a phase angle to the carrier supply so that

$$v=V\sin(\omega_ct+\phi_c) \tag{4}$$

The functions to be used for the display of the LMLT curve are

$$v_N=(a_3)q_N+(b_3)q_N^3 \tag{5}$$

and

$$v_N=(a_5)q_N-(b_5)q_N^3+(c_5)q_N^5 \tag{6}$$

In steady state, the current that flows through C_L must also flow through C_N since $i_c<<i_L$. Therefore

$$i_L\equiv\frac{dq_L}{dt}=\frac{dq_N}{dt} \tag{7}$$

$$\frac{d^2q_L}{dt^2}=\frac{d^2q_N}{dt^2} \tag{8}$$

From equation 7 the time-varying component of q_L is equal to the time-varying component of q_N, hereafter denoted q_{N_T}. The constant component of q_L is $(-C_L V_c)$. Therefore, using the cubic approximation of equation 5, the differential equation can be written

$$L\frac{d^2q_N}{dt^2}+R\frac{dq_N}{dt}+\frac{1}{C_L}q_{N_T}+(a_3)q_N+(b_3)q_N^3 = V\sin(\omega_ct+\phi_c)+V_c \tag{9}$$

With the quintic approximation the equation is

$$L\frac{d^2q_N}{dt^2}+R\frac{dq_N}{dt}+\frac{1}{C_L}q_{N_T}+(a_5)q_N-(b_5)q_N^3+(c_5)q_N^5=V\sin(\omega_ct+\phi_c)+V_c \tag{10}$$

The procedure of solution for either equation 9 or 10 will be exactly the same with the method of solution to be used. For simplicity, therefore, the solution will be developed in terms of equation 9.

Solution of the Differential Equation

In order to introduce additional generality the charge versus potential characteristic of the nonlinear capacitor will be written in terms of the electric gradient E and the electric displacement D for the barium-strontium-titanate dielectric. For a parallel plate capacitor

$$E=\frac{v_N}{l} \tag{11}$$

$$D=\frac{q_N}{A} \tag{12}$$

where l is the thickness and A is the area of the dielectric slab.

Setting

$$(a_1)=\frac{a_3 l}{A} \tag{13}$$

$$(b_1)=\frac{b_3 l}{A^3} \tag{14}$$

equation 5 can be written

$$v_N=\frac{a_3 l}{A}\,q_N+\frac{b_3 l}{A^3}\,q_N{}^3 \tag{15}$$

or

$$E=a_3D+b_3D^3 \tag{16}$$

Changing the dependent variable to electric displacement, equation 9 becomes

$$AL\frac{d^2D}{dt^2}+AR\,\frac{dD}{dt}+\frac{A}{C_L}\,D_T+a_3lD+b_3lD^3 = V\sin(\omega_c t+\phi_c)+V_c \tag{17}$$

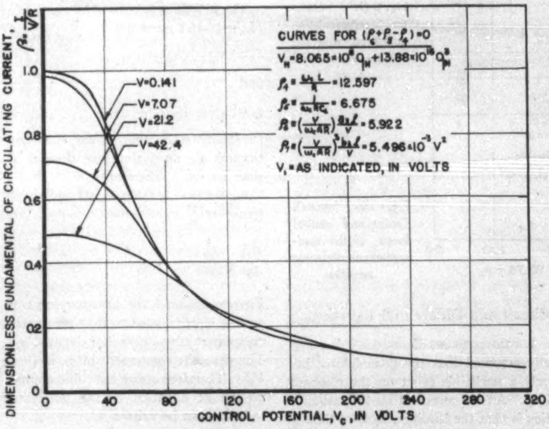


Fig. 10. Tuning curve from the cubic approximation, $(\rho_6+\rho_5-\rho_4)=0$

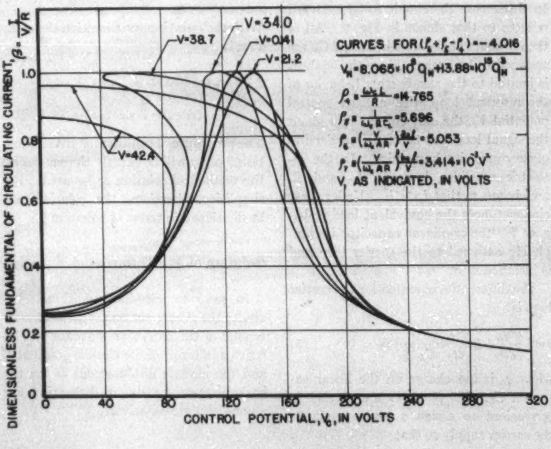


Fig. 11. Tuning curve from the cubic approximation $(\rho_6+\rho_5-\rho_4)=-4.016$

A convenient dimensionless form for equation 17 is

$$\rho_4\frac{d^2\rho_1}{d\rho_2{}^2}+\frac{d\rho_1}{d\rho_2}+\rho_5\rho_{1_T}+\rho_6\rho_1+\rho_7\rho_1{}^3 = \rho_3+\sin(\rho_2+\phi_c) \tag{18}$$

where

$\rho_1=\left(\frac{\omega_c AR}{V}\right)D$ is the new dependent variable. It is the ratio of the instantaneous charge on the nonlinear capacitor to the maximum charge which could appear on it if the circuit were linear and in resonance

$\rho_2=\omega_c t$ is the new independent variable

$\rho_3=\frac{V_c}{V}$ is the ratio of the control to the peak carrier supply potential

$\rho_4=\frac{\omega_c L}{R}$ has the form of an inductive circuit quality factor

$\rho_5=\frac{1}{\omega_c RC_L}$ has the form of a capacitive quality factor

$\rho_6=\left(\frac{V}{\omega_c AR}\right)\frac{a_3 l}{V}$. The quantity $(a$[illegible]$/A$ is the reciprocal of the differential capacitance of the nonlinear capacitor at zero control potential. Note that $\rho_5+\rho_6=[1/(\omega_c R)][1/C_L+(a_3 l)/A]$ would be the total circuit capacitive quality factor if the circuit were linear

$\rho_7=\left(\frac{V}{\omega_c AR}\right)^3\frac{b_3 l}{V}$ includes the nonlinearity

As mentioned earlier, experiments show that the circulating current in the carrier loop remains essentially sinusoidal for conditions of operation encountered in the dielectric amplifier of Fig. 1. In addition the amplifier will not be operated under conditions which result in nonperiodic or subharmonic components of current; the fundamental frequency of the circulating current will be the same as the carrier source frequency. In view of this it seems reasonable that the form of the first approximation of an iterative solution to equation 18 should be

$$\rho_1\cong\alpha+\beta\sin\rho_2 \tag{19}$$

α and β are constants to be determined. Substituting equation 19 into equation 18, ignoring the harmonic frequency terms and equating the coefficients of like variables yields the following relationships

$$\cos\phi_c=\beta\left\{\left[(\rho_5+\rho_6-\rho_4)+\frac{3}{4}\,\rho_7\beta^2\right]+3\rho_7\alpha^2\right\} \tag{20}$$

$$\sin\phi_c=\beta \tag{21}$$

$$\rho_3=\rho_6\alpha+\rho_7\alpha^3+\frac{3}{2}\,\rho_7\alpha\beta^2 \tag{22}$$

Squaring equations 20 and 21 and adding

Vol. 7
Pt. 1
1955

$$\beta^2\left\{1+\left[3\rho_7\alpha^2+(\rho_6+\rho_5-\rho_4)+\frac{3}{4}\,\rho_7\beta^2\right]^2\right\}$$
$$=\sin^2\phi_c+\cos^2\phi_c=1 \qquad (23)$$

Solving equation 23 for α^2

$$\alpha^2=\frac{1}{3\rho_7}\left\{\pm\sqrt{\frac{1}{\beta^2}-1}-\left[(\rho_6+\rho_5-\rho_4)+\frac{3}{4}\,\rho_7\beta^2\right]\right\} \qquad (24)$$

Equations 22 and 24 constitute two independent relations for the two unknowns, and it is possible to find α and β from them.

As stated at the outset, it is desired to find the tuning curves for the resonant-circuit dielectric amplifier. These curves relate the amplitude of the circulating current in the resonant carrier loop to the control potential applied to the nonlinear capacitor. Since ρ_1 has the form of a dimensionless charge, $(d\rho_1)/(d\rho_2)$ must have the form of a dimensionless current. From the definitions and equation 7

$$\frac{d\rho_1}{d\rho_2}=\frac{\dfrac{dq_N}{dt}}{\dfrac{V}{R}}=\frac{i_L}{\dfrac{V}{R}} \qquad (25)$$

From the first approximation solution

$$\frac{d\rho_1}{d\rho_2}\cong\beta\cos\rho_2 \qquad (26)$$

In view of equations 25 and 26

$$i_L\cong\beta\frac{V}{R}\cos\omega_c t \qquad (27)$$

Apparently β is the ratio of the amplitude of the fundamental component of the circulating current to the amplitude of the resonant current which would flow if the circuit were linear and in resonance. Also, from the definitions

$$V_c=\rho_3 V \qquad (28)$$

Hence, a curve of β as a function of ρ_3 or V_c for fixed values of the remaining parameters will have the form of the aforementioned tuning curve.

The elimination of α or β between equations 22 and 24 results in a complicated function implicitly relating the unknown and the parameters. Fortunately it is not necessary to attempt the formal solution of these equations. Instead, the following method of solution by computation is available.

1. A numerical value is selected for β.
2. The corresponding α is obtained from equation 24.
3. ρ_3 (and hence V_c) for this α and β is then given by equation 22.

An approximate solution can also be developed for the differential equation which includes the quintic term in the nonlinear charge versus voltage relation. Proceeding in the manner outlined, for the cubic approximation again yields a pair of equations relating α, β, and the parameters. The procedure of computation used in plotting β as a function of ρ_3 or V_c can be the same as that selected for the results of the cubic approximation.

Analytical and Experimental Curves

The form of the tuning curve most convenient for computation of amplifier gain is a plot of β versus V_c. Figs. 10 and 11 show tuning curves obtained using the cubic approximation to the capacitor nonlinearity. Only positive values of V_c are shown since there is symmetry about the β axis. On each figure one curve is plotted for the case where the carrier potential swing across the nonlinear capacitor is such that the circuit is very nearly linear operation. Two or more curves are included on each graph for much larger values of excitation. The "distortion" of the steady-state tuning characteristics due to the nonlinearity is apparent from these curves.

The jump phenomenon, often observed in the second-order nonlinear circuit as the carrier frequency is varied, may also occur in the carrier circuit of Fig. 1 as the

Fig. 12 (right). Tuning curve from the quintic approximation

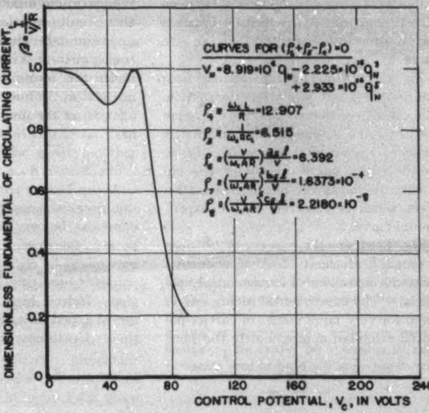
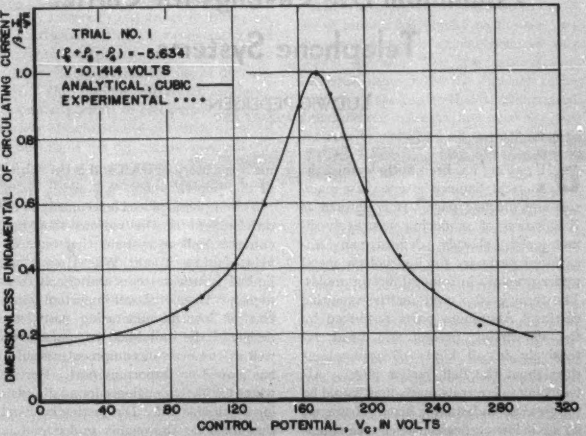


Fig. 13 (below). Comparison of an analytical curve and experimental points for a small value of carrier supply potential



ol. 74
. 1
55

control potential is varied. On the curve for $V=38.7$ volts in Fig. 11 there is a region, between $V_c=40$ and $V_c=96$ volts, where three values of β correspond to a single V_c. If the circuit is initially in steady-state operation at some V_c greater than 96 volts and the control potential is very slowly decreased, the circuit will operate along the top portion of the curve until $V_c=40$ is reached. Then the circulating current will suddenly jump to the smaller β for $V_c=40$ volts. If, now the control potential is slowly increased from $V_c=40$, the circuit will be in operation on the lower portion of the curve until $V_c=96$ volts is reached at which time the current will jump back up to the higher β. The circuit will not assume stable operation for those β's between $V_c=40$ and 96 volts, which lie between the upper and lower portions of the curve.

Fig. 12 is a tuning curve obtained from the results of the quintic approximation. The circuit parameters for this curve are not the same as the circuit parameters for the experimental curve of Fig. 6. However, the curve does show the dip in current due to the initial polarization effect which is observed on the experimental curve.

Fig. 13 shows the agreement between an analytical curve and experimental points for one value of carrier supply potential. The experimental tuning curves taken for very large values of carrier potential exhibited approximately the same shape as the corresponding analytical curves. However, in some cases the analytical and experimental curves failed to show accurate point-by-point agreement. An investigation indicated that this discrepancy was due primarily to the difference between the LMLT curve and the actual minor loop locus when large carrier swings appeared across the nonlinear capacitor.

Voltage Gain

The low-signal frequency voltage gain of the resonant-circuit dielectric amplifier can be obtained directly from the tuning curve for the circuit. To avoid distortion of the signal, the direct bias voltage and the maximum input signal are chosen so that circuit operation takes place over an approximately straight portion of the tuning curve. With this restriction, and under the assumption of an ideal demodulator, the low-signal frequency voltage gain of the amplifier shown in Fig. 1 is[1]

$$VA = GQV \qquad (29)$$

As can be seen from Fig. 11, under certain circumstances the slope of the tuning curve can be very large. It is to be expected, therefore, that a large voltage gain might be obtained from a dielectric amplifier operating in the high slope region. Indeed, high gains have been observed experimentally for short periods of time. Up to now, however, such operation has been unstable and frequent retuning of the circuit has been required in order to maintain the amplification. Part of the instability can probably be attributed to inadequate control of the sample temperature. Part may be due to the unstable behavior of the material itself. Additional experimental work is in progress at the Carnegie Institute of Technology to determine whether circuit operation over the high slope portions of the tuning characteristics is feasible.

References

1. **Dielectric Amplifiers, G. W. Penney, J. R. Horsch, E. A. Sack.** *AIEE Transactions*, vol. 72, pt. I, Mar. 1953, pp. 68–79.

2. **Frequency Response of a Resonant Dielectric Amplifier, G. W. Penney, E. A. Sack, E. R. Wingrove.** *Ibid.*, vol. 73, pt. I, May 1954, pp. 118–24.

3. **A Mathematical Analysis of the Dielectric Amplifier, L. A. Pipes.** *Journal of Applied Physics*, New York, N. Y., vol. 23, no. 8, Aug. 1952, pp. 818–24.

4. **A Mathematical Analysis of a Series Circuit Containing a Nonlinear Capacitor, Louis A. Pipes.** *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 238–44.

5. **Graphical Analysis of Nonlinear Circuits Using Impedance Concepts, J. S. Thomsen.** *Journal of Applied Physics*, New York, N. Y., vol. 24, no. 11, Nov. 1953, p. 1379.

6. **Analytical Calculation of Circuits with Nonlinear Capacitors, D. M. Kazarnovski.** *Elektrichestvo*, Moscow, USSR, no. 10, Oct. 1953, p. 68.

7. **Forced Oscillations of Nonlinear Circuits, Louis A. Pipes.** *AIEE Transactions*, vol. 73, pt. I, Sept. 1954, pp. 352–58.

8. **Variation with Time of the Dielectrical Properties of Ferroelectric Ceramics, D. M. Kazarnovski.** *Zhurnal Teknicheskoi Fisiki*, Moscow, USSR, vol. 22, Apr. 1952, p. 553.

Aluminum Die Castings for Carrier Telephone Systems

LUDWIG PEDERSEN
ASSOCIATE MEMBER AIEE

DIE casting has been aptly termed the shortest distance between raw material and finished part. It is defined as the process of producing accurately dimensioned, sharply defined, smooth-surfaced parts by forcing molten metal under pressure into metal dies or molds. The term is also used for the resulting casting. Aluminum parts produced by the die-casting process are used extensively in all kinds of applications throughout the Bell System plant. Attempting to cover its many uses would be too great an undertaking here. However, its use in low-cost carrier telephone equipment is a story in itself and is the subject of this paper.

The emphasis placed on economy in the development of the various short-haul carrier telephone systems that have developed since World War II was not limited to new systems and circuit techniques. To meet the all-important objective of low manufacturing costs, the design of the individual components as well as the over-all equipment assembly has played an important part. Here is where the design engineer found die casting a valuable aid. Die casting is an art that must be thoroughly understood in order that the designer can realize a substantial part of its possibilities, which are practically unlimited in scope. The uniformity of die castings shows to great advantage on the assembly line of high-production equipment. The time-consuming annoyances of misalignments and necessary adjustments usually present with fabricated chassis are nonexistent with properly designed castings. Die castings can be machined readily when this is required, and are uniform in dimensions, thus facilitating assembly and aiding the interchangeability of parts. Temperature variations encountered in

---

Paper 55-259, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Electrical Utilization of Aluminum Conference, Pittsburgh, Pa., March 15–17, 1955, and re-presented at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted December 17, 1954; made available for printing January 27, 1955.

LUDWIG PEDERSEN is with the Bell Telephone Laboratories, Inc., Murray Hill, N. J.

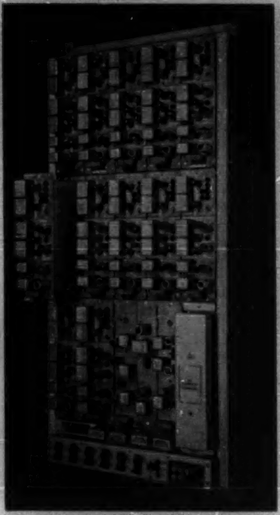


Fig. 1. A fully equipped 12-channel N1 carrier terminal. Channel unit 6 shown partly pulled out from supporting frame

the use of carrier equipment do not appear to cause any permanent dimensional changes or loss of physical properties.

Aluminum was chosen for this application because of its light weight (approximately 40 per cent of that of zinc die castings), mechanical strength, low cost, dimensional stability, good electrical shielding properties in the carrier frequency range, and the rather pleasing appearance of the natural aluminum finish. The aluminum alloy used in these die castings contains 87.5 per cent aluminum, 9 per cent silicon, and 3.5 per cent copper. It has a tensile strength of 40,000 pounds per square inch, a yield strength of 25,000 pounds per square inch, and an impact strength of 3 foot-pounds.

Aluminum alloys form and maintain a thin transparent oxide film which, for many uses, is sufficient protection for the die casting, and no further finishing is necessary. This is an item of importance when cost is a design consideration. A variety of finishes, however, may be applied. These may consist of organic, electrolytic, oxide, electroplating, chemical, or mechanical treatments.

Many questions usually confront the design engineer when he attempts to decide between the use of die castings and fabricated structures. The all-important question of cost is primarily one of comparison between unit cost of a die casting versus a fabricated structure. Large demands are not necessarily the criterion for justifying a die casting economically; complexity of design is also an important factor. A simple shape that can be punched or extruded would not, as a rule, be a good candidate for the die-casting technique. It may sometimes prove advisable to use another method such as sand or permanent-mold casting, particularly for low-demand items. However, it is well to remember that because of the fine-grain structure obtained by the rapid chilling of the alloy in the die, die castings have an "as-cast" strength that is greater than that of castings made by any other method provided, of course, they are made of the same alloy. As a rough guide to the designer, however, it may be stated that an item of fair complexity and with an annual demand of around 5,000 would justify die casting. Usually the die cost is found to be relatively high, but the unit cost low. Consideration must also be given to other factors besides cost, such as uniformity and appearance.

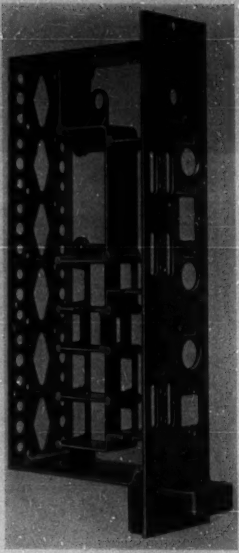


Fig. 2. Die casting for one of subassemblies common to N1 carrier terminal. On front of casting, equipment designations incorporated in die. Large number of cored holes in rear surface provides clearance for filter terminals as well as filter-mounting holes. In middle portion of die casting are a number of compartments used to hold miniature transformers

The *N1* carrier telephone system[1] was the first of the developments to employ die castings. A 12-channel terminal as shown in Fig. 1 uses 39 aluminum die castings involving six individual designs or dies. Size and weight of the equipment was minimized by arranging the miniaturized components compactly in plug-in units using die-cast aluminum chassis. These units are of a shape and size to utilize fully the relay rack space available in depth as well as in breadth and height. It is in this type of design, usually referred to as cubic construction, as contrasted with planar construction of conventional panels, that the design engineer benefits from the use of die castings. The various mounting surfaces, as well as pockets or compartments for a variety of components, are easily obtained by means of the die-casting process. In the case of the *N1* carrier development, the decision to proceed with aluminum die castings was based on extensive comparison with other methods of fabrication. The number of castings used in each system varies from two to as high as 24 castings of one kind. The die casting shown in Fig. 2 is a small-demand item numbering about 1,000 units a year. Even with this relatively small demand it was found that the unit cost (including some machinery which consisted of drilling and tapping a number of holes) and its share of the die cost amortized over a 4-year period was found to be only half the cost of a fabricated unit.

The three die castings shown in Fig. 3 are of similar complexity to the one just discussed but the annual demand was 24,000 of each. The cost estimates in this case showed the die castings to be only about one-fourth that of a fabricated unit. These castings are used for the *N1* carrier channel unit, which consists of three separate subassemblies mounted together to form a complete plug-in unit. Fig. 4 shows the three castings assembled as they appear in the channel unit. The unit, fully equipped, is shown in Fig. 5.

In addition to the differences in the cost of producing these castings, there are other cost differences favoring the die castings. If properly designed, the die castings will incorporate details that otherwise need to be mechanically fastened to the fabricated framework, and such items as equipment designations used to identify components may be incorporated in the die by the use of raised characters in a recessed area, instead of

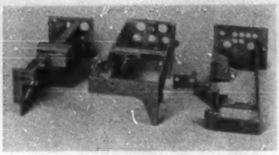


Fig. 3 (left). Rear view of three die castings used for N1 carrier channel unit. Two castings to right were later used for O1 carrier channel unit as well. All large holes shown in front surfaces are cored

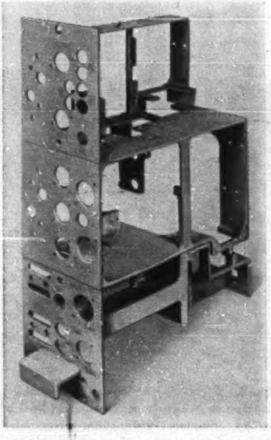


Fig. 4 (left). Three die castings of Fig. 3 assembled as they appear in N1 carrier channel unit. In lower portion of assembly may be seen horizontal partitions cast in to provide required electrical shielding

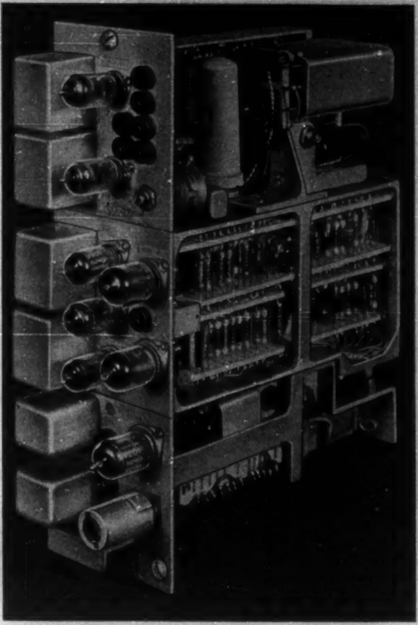


Fig. 5 (right). Front view of N1 carrier channel unit fully equipped. In service, unit is equipped with aluminum can cover not shown

Vol
Pt.
195

by being applied by various printing methods.

In evaluating some of the experiences gained during the design period of the *N1* carrier equipment, it is interesting to note that the first lesson to be learned for the design engineer is to shape his thinking in terms of die casting. This ability comes only with experience and is absolutely necessary to take the best advantage of the die-casting art. The designer should be able to build into the casting the many possible advantageous features, and the art provides him with wide latitude to accomplish this objective. In other words, the part must be designed specifically for this method of manufacture. In the early design stages, the closest co-operation is necessary between the design engineer and die designer. The electrical requirements, together with the size, shape and functions of all components should be furnished to the die designer so that he may assist in reaching the design goals through skillful application of die casting. Another lesson is that unnecessary accuracy means added tool expense. It is good engineering to limit the use of close tolerances to the really important dimensions. In some instances such tolerance can be obtained more economically by subsequent machining of the part. It is not advisable to attempt to hold accurate dimensions across the parting line. Also the design engineer must dimension from a common reference, rather than build one dimension upon another. The minimum tolerances specified for the carrier die castings discussed here were held to ±0.004 inch on any dimension, although ±0.002 is a tolerance common in die-casting practice.

Concerning wall thickness, it is good practice to use thin walls and obtain required strength by the use of ribs. There is, of course, a practical minimum to the thickness of wall specified. One governing factor is the necessity of filling the die properly. The minimum wall thickness specified for the castings referred to here is 7/64 inch. It is desirable for wall sections to be of uniform thickness for lowest production cost. Where variations in thickness are necessary, the transition from thin to heavy sections should be as gradual as possible. However, it is well to remember that each design should be treated as an individual problem. This important rule applies equally to such items as size and weight of a casting, tapers, and hole diameters. The word "taper" is one to be treated with considerable respect since the die designer is constantly concerned with its implications. "Taper or draft" is the term for the slopes of the walls of sections and sides of holes to facilitate withdrawing the casting from the die without tearing, sticking and distortion.

It is interesting to note the progress in the use of die castings as the development of the *01* carrier system proceeded.[3] Basically, the same design approach was employed here as was used in the *N1* carrier system. In fact, two of the die castings used in the *N1* carrier channel unit (two right-hand units shown in Fig. 3) were specified without any changes whatsoever. In the *N1* carrier, the terminal equipment with its large number of plug-in units using die-cast chassis, employed a supporting framework fabricated from sheet aluminum. However, in the *01* carrier die-cast shelves mounted directly on the relay rack are used. Two identical shelves are used in each terminal mounting as shown in Fig. 6. Note that the lower shelf is turned upside down. When the plug-in

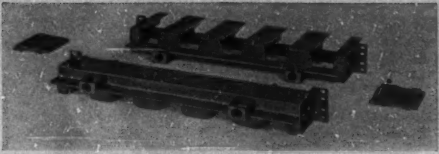


Fig. 6 (above). Die-cast shelf assembly used in O1 carrier terminal. Two identical shelf castings are used for each terminal mounting. Lower shelf is shown upside down, which is position it takes in final terminal mounting assembly. Shelves are 19 inches long and arranged to mount directly on standard relay racks

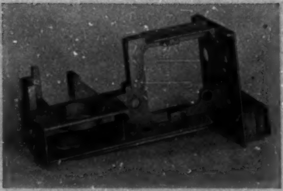


Fig. 9 (right). Die casting designed to form part of O1 carrier channel unit provides excellent chassis for variety of components and incorporates a number of mechanical features such as guide arrangement referred to in Fig. 8. At right, handle located on front surface. Die for casting required four slides to provide for all necessary features

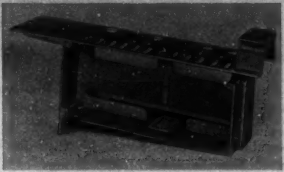


Fig. 10. Design where die required four slides. Note thin wall sections

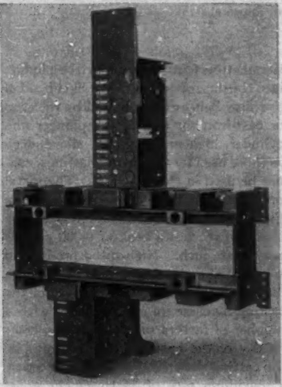


Fig. 7. Two die-cast shelves comprising O1 carrier terminal mounting with die casting for one of plug-in units in position, resting on top shelf, with another die-cast chassis suspended from lower shelf. Five plug-in units supported from each shelf

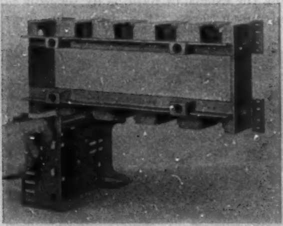


Fig. 8. Slide arrangement which guides plug-in units into place

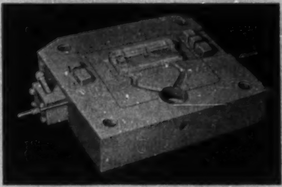


Fig. 11. Cover half of die used to produce die casting shown in Fig. 2. Set-in impression block in place

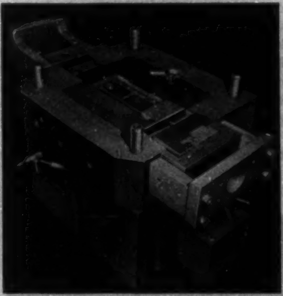


Fig. 12. Die base assembly with ejector die-half. Main part of die with angular members, side-pulled slides, and main impression block in place

units are inserted, those on the top rest on the shelf while those on the bottom are suspended from the shelf as shown in Fig. 7. Fig. 8 shows the slide arrangement. The slides for guiding the units into position are part of the unit die castings and this feature serves the important function of accurately mating the plug and socket required for the electric connections between the plug-in unit and the terminal mounting proper. In general, the die castings designed for the *01* carrier incorporated more desirable features than the earlier designs for the *N1* carrier. The die casting shown in Fig. 9 is an exceptionally fine example of a complex design which was a challenge not only to the carrier equipment designer, but also to the die designer. Fig. 10 shows another example.

The die casting shown in Fig. 2 is used in the *N1* carrier and is typical of the complex structures that can be made by the die-casting process. Figs. 11 and 12 show two views of the large steel die used in producing this die casting. When ready for the casting machine, the die weighs about 3,000 pounds and consists of a number of precisely made parts expertly fitted together. It is made of two matching halves of polished steel cut to form a cavity between them shaped exactly like the final part desired. The two halves are brought together in the die-casting machine and the molten metal is injected under pressure into the cavity between the two halves. After the metal solidifies, the machine opens the die, and the completed casting is ejected. Fig. 11 illustrates the cover half of the die and contains an impression block; Fig. 12 shows the die-base assembly with the ejector die-half in place. This half contains the main parts of the die, including the movable cores or slides for making the holes or slots in the casting. The dies may have a single cavity for making one casting at each cycle of the machine, as in the die just described. However, when comparatively small parts of suitable shape are to be cast in large quantity, several cavities may be incorporated in the die. If such cavities are duplicates, the die is called a multiple-cavity die. When the cavities form parts of varying shape, the die is known as a

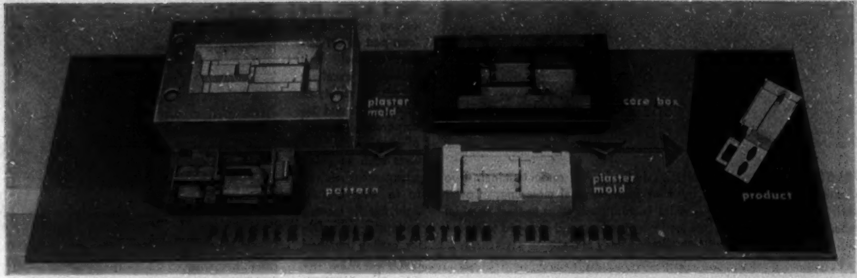


Fig. 13. Display of plaster molds, wooden core box, and pattern for part shown at right

combination die. Some of the dies used for the carrier die castings can be employed for several castings simply by changing one or more of the movable cores or slides. The one illustrated is such a die, having two slides available for producing two different front surfaces, leaving all other parts of the die the same. Obviously, this feature reduces tool costs an appreciable amount. The tooling interval that was involved for this rather complex die was 14 weeks. A good point for the designer to keep in mind is that the die maker is the highly skilled mechanic who must make the die in reverse.

Since the greater percentage of the die cost is in labor, it is important that the steel be adaptable to machine processes. Fully annealed high-quality steel is used, with full heat treatment following the completion of the die. From here on, there are a number of problems such as heat-treating characteristics, heat checking, cracking, distortion, and resistance to deformation, which are the concern of the designer and not within the scope of this paper.

The total number of castings that may be produced from one die depends upon the type of castings. The dies used for the carrier castings are expected to produce up to 200,000 castings each. Design changes affecting the completed die are not desirable, but if made are usually expensive as well as risky, since the die may be distorted to the extent that an entirely new die must be made.

Two basic classifications of die-casting machines are in use today: 1. the "submerged-plunger" type, used for casting tin, lead, and zinc, and 2. the "cold-chamber" machine, used for aluminum, magnesium, and copper. The aluminum alloy castings for the carrier systems are produced on a cold-chamber machine, which type of machine requires that prior to each casting the molten metal be ladled, either manually or automatically, into the injection cylinder. From there it is forced into the die by a hydraulically operated plunger. On the other hand, in the submerged-plunger type of machine the plunger and cylinder are always submerged in the molten metal in the furnace. At the end of each "shot," metal flows into the cylinder, and at the start of the next cycle it is forced by the plunger into the die. Because of the likelihood of the alloy being contaminated by the steel or cast-iron surfaces of the plunger and cylinder, high melting-point alloys, such as those used for the carrier castings, cannot be used in this machine. The metal pressure used for the casting shown in Fig. 2 starts at 4,500 pounds per square inch and reaches a peak of 9,000 pounds per square inch. High-speed production is one of the principal advantages of the die-casting process. The capacity of the casting machine, of course, varies with a number of factors such as size and complexity of the casting. As an example, the casting shown in Fig. 2 is produced at a maximum rate of 75 to 77 per hour. With small, fairly simple parts several hundred shots per hour is not unusual. Even with large complex units, the casting cycle is relatively short.

The casting cycle having been completed, the casting will require removal of excess metal. The usual way of removing the excess, or flash, is by shearing in a punch press. For items in large

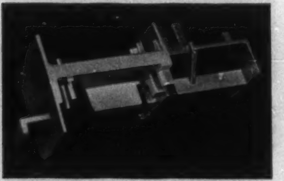


Fig. 14. Aluminum casting made from plaster mold, cast with tolerance of ±0.005. A 125-microinch finish obtainable from plaster-molded castings similar in appearance even though not quite as smooth as die castings

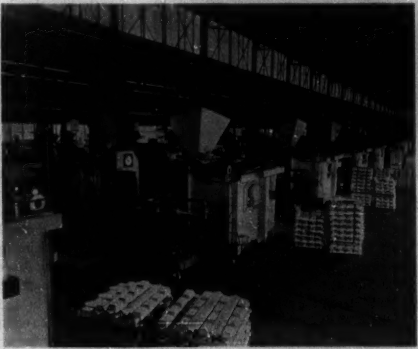


Fig. 15. Modern aluminum die-casting plant showing individual melting furnaces and tracks for quickly moving machines. Machines moved to service shop for changing dies and for maintenance with only a few minutes lost time for position

Courtesy Doehler-Jarvis Division of National Lead Company demand special jigs are constructed for this purpose. In a number of instances hand filing serves the purpose.

Machining of the castings, where required, is usually limited to drilling of holes too small to be cored economically (less than 1/8 inch in diameter), tapping, and squaring surfaces where taper required in casting cannot be tolerated. Where the dimensional accuracy required is such that it cannot readily be provided by the die, a machining operation may be indicated. The necessary machining is usually the business of the die caster, and no further work is required on the casting by the user. Since the machining adds to the parts cost, it is good engineering to keep it to a minimum. One way of doing this is to be sure that, if screws are used as fastening devices for the components, they are large enough so that holes may be cored rather than drilled. The use of self-tapping screws avoids the tapping operation, etc.

Before production of carrier systems can begin, it is usually necessary to conduct extensive field trials of the equipment. This requires models that must be very similar to the final product. The location of each component as well as the path of each wire must be the same in both cases; otherwise, the field trial results are of little or no value. Having decided that production will use aluminum die casting with all its built-in advantages, the problem is how to construct the models, a process which, until recently, was both expensive and time-consuming. Sheet aluminum was used in fabricating the models which had to simulate the final die-cast product. The drafting effort alone was considerable, and once the models were built, the drawings were of no further value. Today, the carrier models are produced by means of plaster molds. The production drawings for the final die casting are used to construct wooden patterns from which the plaster mold is produced. This is a quick and relatively inexpensive way of obtaining a small number of models and also serves as a check on the die-casting design. Fig. 13 shows a display of a plaster mold and the wooden pattern. A casting made from this plaster mold is shown in Fig. 14. Fig. 15 shows a modern aluminum die-casting plant.

References

1. A NEW TELEPHONE CARRIER SYSTEM FOR MEDIUM-HAUL CIRCUITS, R. S. Caruthers, H. R. Huntley, W. E. Kahl, Ludwig Pedersen. *Electrical Engineering*, vol. 70, Aug. 1951, pp. 692–7.

2. TYPE-O CARRIER TELEPHONE, J. A. Coy, E. K. Van Tassel. *AIEE Transactions*, vol. 71, pt. I, 1952 (Jan. 1953 Section), pp. 428–37.

---

ol, 74
. 1
955

Analysis of Nonlinear Coupled Circuits—Part II

Y. H. KU
FELLOW AIEE

IN A paper presented in 1954,[1] the analysis of nonlinear coupled circuits was made by the method of high order phase planes or phase space. The simultaneous differential equations are first combined into one higher order differential equation and then solved by a method outlined in the previous paper and illustrated by two examples. However, when there are nonlinearities existing in both loops of, say, a 2-loop coupled circuit, it would be difficult to apply the procedure given previously. An alternative is here presented by working with the set of differential equations as originally given and solving several phase plane equations simultaneously. The new method may be called the method of solving simultaneous phase plane equations, in contrast with the method of solving one higher order phase space equation. This method has the advantage that nonlinearities may exist in two or more original differential equations. In a 2-loop circuit for example, there may be a nonlinear resistor in the primary loop and a nonlinear inductor in the secondary loop. Using examples I and II in the previous paper, the relations between the primary and secondary currents and between their first derivatives are shown. Example III gives the results of phase plane trajectories for a circuit with nonlinearities in both the primary and the secondary loops. Example IV shows that for a coupled circuit with a nonlinear capacitor in the secondary loop (the nonlinear capacitance being a function of voltage) the primary and secondary currents can be correlated by one equation in the phase space and solved simultaneously.

General Outline of Method

The method can be explained in connection with a 2-loop coupled circuit with a resistor, an inductor, and a capacitor in each loop and coupled by a mutual inductance; see Fig. 1. The primary resistance has a constant part and a nonlinear part (as a function of the primary current). The secondary inductance has a nonlinearity which is a function of the secondary current. The differential equations are

$$R_1(1+\rho)i_1+(L_1+M)pi_1+q_1/C_1-Mpi_2=e_1 \quad (1)$$

$$R_2i_2+(L_2\beta+M)pi_2+q_2/C_2-Mpi_1=0 \quad (2)$$

where i_1 and i_2 represent the loop currents; q_1 and q_2 represent the loop charges; e_1 represents the applied voltage in loop 1; R_1 and R_2 represent resistances; L_1 and L_2 represent self-inductances; M represents coupling inductance; and C_1 and C_2 represent capacitances. The symbol p stands for the time differential operator d/dt; ρ as a function of i_1 signifies the nonlinear part of the primary resistance; and β as a function of i_2 signifies the nonlinear element in the secondary loop inductance. Let

$$q_1=x_1,\ i_1=pq_1=v_1,\ pi_1=p^2q_1=g_1$$

$$q_2=x_2,\ i_2=pq_2=v_2,\ pi_2=p^2q_2=g_2$$

The general outline of the method is comprised of the following steps:

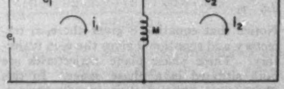


Fig. 1. Coupled circuit

Paper 55-428, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted October 14, 1954; made available for printing April 1, 1955.

Y. H. Ku is with the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.

Example IV was worked out by Murray Falkowitz in a graduate course on nonlinear analysis given by the author in 1953–54.

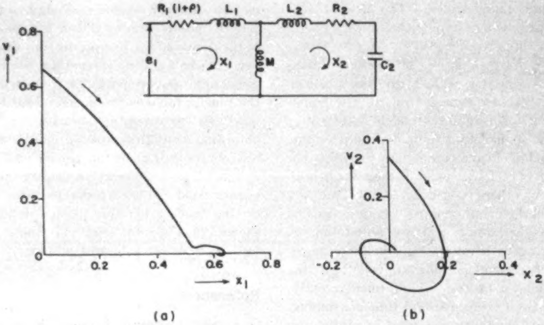


Fig. 2. Trajectories for example I

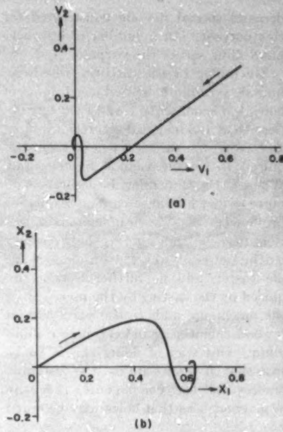


Fig. 3. Trajectories for example I

1. Rewrite equations 1 and 2, using the symbols x_1, x_2, etc.

$$(L_1+M)g_1 - Mg_2 = e_1 - R_1(1+\rho)v_1 - x_1/C_1 = e_1' \quad (3)$$

$$-Mg_1 + (L_2\beta+M)g_2 = -R_2v_2 - x_2/C_2 = e_2' \quad (4)$$

2. Solve for g_1 and g_2 (the first derivatives of i_1 and i_2).

$$g_1 = \frac{(L_2\beta+M)e_1' + Me_2'}{D} \quad (5)$$

$$g_2 = \frac{(L_1+M)e_2' + Me_1'}{D} \quad (6)$$

where

$$\begin{aligned} D &= (L_1+M)(L_2\beta+M) - M^2 \\ &= (L_2\beta+M)L_1 + ML_2\beta \\ &= (L_1+M)L_2\beta + ML_1 \end{aligned} \quad (7)$$

3. Write two-phase plane equations

$$\frac{dv_1}{dx_1} = \frac{g_1}{v_1} \quad (8)$$

$$\frac{dv_2}{dx_2} = \frac{g_2}{v_2} \quad (9)$$

where g_1 and g_2 are given by equations 5 and 6 for any set of values of e_1, x_1, x_2, v_1 and v_2. Notice that $g_1 = pv_1$, $g_2 = pv_2$, $v_1 = px_1$ and $v_2 = px_2$.

4. Correlate the change of v_1 with the change of v_2 by the first order differential equation in phase space

$$\frac{dv_2}{dv_1} = \frac{g_2}{g_1} \quad (10)$$

Similarly, the change of x_1 can be correlated with the change of x_2 by the first-order differential equation in phase space

$$\frac{dx_2}{dx_1} = \frac{v_2}{v_1} \quad (11)$$

Notice that equation 8 gives the v_1-x_1 trajectory and equation 9 gives the v_2-x_2 trajectory. These phase plane trajectories are both situated in a phase space. In this phase space, the v_2-v_1 trajectory (in the v_2-v_1 plane) and the x_2-x_1 trajectory (in the x_2-x_1 plane) may also be obtained.

5. Given the initial values $x_1 = q_1$, $x_2 = q_2$, $v_1 = i_1$ and $v_2 = i_2$, calculate the initial values of g_1 and g_2 according to equations 5 and 6. Calculate also the initial slopes of the v-x trajectories according to equations 8 and 9.

6. Starting from the initial values of x_1 (at $t = 0$), assume a small change Δx_1. From the initial slope given in equation 8, calculate Δv_1. Obtain a new set of values for x_1 and v_1 at the end of the interval by adding these increments to the initial values. By taking the average of v_1 in the beginning of the interval and at the end of the interval, $\Delta x_1 = v_{1av}\Delta t$ is obtained, and the actual time interval Δt can be calculated.

7. According to equations 10 and 11, calculate the increments Δv_2 and Δx_2 corresponding to the increments Δv_1 and Δx_1. Check by the relation $\Delta x_2 = v_{2av}\Delta t$, where Δt is obtained in step 6.

8. With the new set of values of x_1, x_2, v_1, and v_2 (at the end of the chosen interval or, more accurately, at the middle of the interval), calculate g_1 and g_2 according to equations 5 and 6. Also calculate the new slopes dv_1/dx_1 and dv_2/dx_2 according to equations 8 and 9.

9. Repeat steps 6, 7, and 8 to obtain new sets of values of x_1, x_2, v_1, and v_2, and also

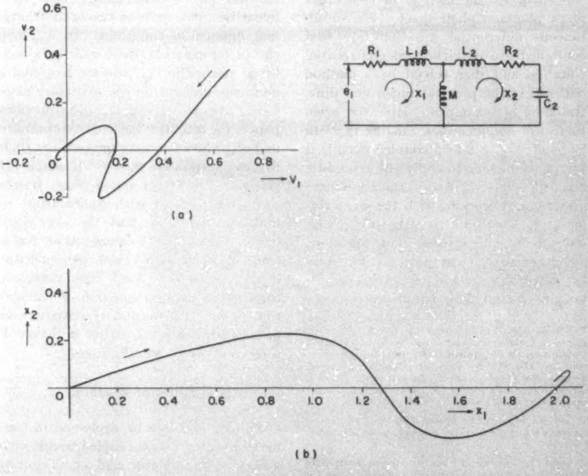


Fig. 4. Trajectories for example II

Vol.
Pt.
195

new values of slopes. Plot v_1-x_1 and v_2-x_2 trajectories and also x_1-t and x_2-t curves.

EXAMPLE I

Example I corresponds to Example I of the previous paper.[1] The differential equations are

$$R_1(1+i_1)i_1+(L_1+M)pi_1-Mpi_2=e_1 \qquad (12)$$

$$R_2i_2+(L_2+M)pi_2+q_2/C_2-Mpi_1=0 \qquad (13)$$

Differentiating with respect to time, equations 12 and 13 become

$$R_1(1+2x_1)v_1+(L_1+M)g_1-Mg_2 = pe_1 \qquad (14)$$

$$R_2v_2+(L_2+M)g_2+x_2/C_2-Mg_1=0 \qquad (15)$$

where $x_1=i_1$, $x_2=i_2$, $v_1=pi_1$, $v_2=pi_2$, $g_1=p^2i_1$ and $g_2=p^2i_2$. (Notice that x corresponds to currents instead of charges.) For $e_1=1.0=\text{constant}$, $pe_1=0$.

The numerical values are $R_1=R_2=1$, $L_1=L_2=M=1$, $C_2=1$ and $e_1=1$. The initial values are $x_1=0$, $x_2=0$, $q_2=0$, $v_1=2/3$, $v_2=1/3$, $g_1=-5/9$ and $g_2=-4/9$. The initial slopes are $dv_1/dx_1=-5/6$ and $dv_2/dx_2=-4/3$.

Rewriting equation 12 for the numerical example gives

$$v_2=(1+x_1)x_1+2v_1-1$$

So for every new set of values of x_1 and v_1, there is a corresponding value of v_2. The new value of x_2 can then be obtained from the relation given in equation 11. That is, for a given increment Δx_1, the corresponding increment Δx_2 is given by $\Delta x_2=(v_2/v_1)\Delta x_1$. (Since both the old set and the new set of values of v_1 and v_2 are known, the slope dx_2/dx_1 can be accurately determined.)

The v_1-x_1 and v_2-x_2 trajectories are shown in Fig. 2, and the v_2-v_1 and x_2-x_1 trajectories in Fig. 3.

EXAMPLE II

Example II corresponds to example II of the previous paper. The differential equations are

$$R_1x_1+(L_1\beta+M)v_1-Mv_2=e_1 \qquad (16)$$

$$R_2x_2+(L_2+M)v_2+q_2/C_2-Mv_1=0 \qquad (17)$$

where $x_1=i_1$, $x_2=i_2$, $v_1=pi_1$ and $v_2=pi_2$. The factor β is a function of $x_1=i_1$ and is taken as equal to $1/\cosh^2 x_1$ in this example.

The numerical values are $R_1=R_2=0.5$, $L_1=L_2=M=1$, $C_2=1$ and $e_1=1$. The initial values are $x_1=0$, $x_2=0$, $q_2=0$, $v_1=2/3$, $v_2=1/3$, $g_1=-5/18$ and $g_2=-2/9$. The initial slopes are $dv_1/dx_1=-5/12$ and $dv_2/dx_2=-2/3$.

The v_1-x_1 and v_2-x_2 trajectories are given in the previous paper. The v_2-v_1 and x_2-x_1 trajectories are illustrated in Fig. 4.

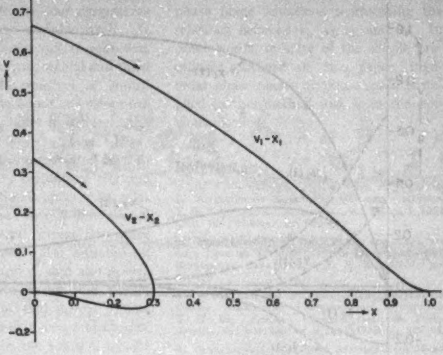


Fig. 5 (right). Trajectories for example III

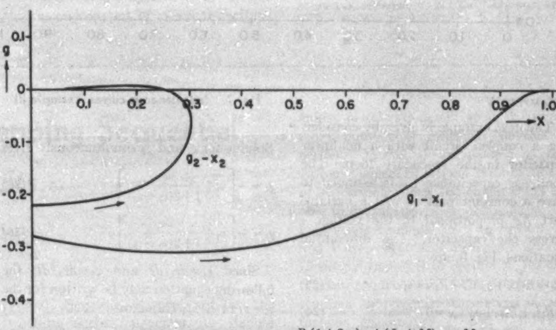


Fig. 6 (below). Trajectories for example III

EXAMPLE III

Referring to Fig. 1 and omitting the capacitors gives a 2-loop circuit with a nonlinear resistor in the primary loop and a nonlinear inductor in the secondary loop. The differential equations are

$$R_1(1+\rho)i_1+(L_1+M)pi_1-Mpi_2=e_1 \qquad (18)$$

$$R_2i_2+(L_2\beta+M)pi_2-Mpi_1=0 \qquad (19)$$

where $\rho=i_1$ and $\beta=1/\cosh^2 i_2$ in this example.

Denoting i_1 and i_2 by x_1 and x_2 and (pi_1) and (pi_2) by v_1 and v_2 respectively results in

$$R_1(1+x_1)x_1+(L_1+M)v_1-Mv_2=e_1 \qquad (20)$$

$$R_2x_2+(L_2\beta+M)v_2-Mv_1=0 \qquad (21)$$

The numerical values are $R_1=R_2=0.5$, $L_1=L_2=M=1$ and $e_1=1$. The initial currents are $x_1=i_1=0$ and $x_2=i_2=0$. Since at $t=0$, $\beta=1$, equations 20 and 21 give the initial values $v_1=2/3$ and $v_2=1/3$.

Differentiating with respect to time, equations 20 and 21 become

$$R_1(1+2x_1)v_1+(L_1+M)g_1-Mg_2 = pe_1 \qquad (22)$$

$$R_2v_2-2\alpha\beta L_2v_2^2+(L_2\beta+M)g_2-Mg_1=0 \qquad (23)$$

where $g_1=pv_1=p^2x_1=p^2i_1$, $g_2=pv_2=p^2x_2=p^2i_2$, $\alpha=\tanh x_2$. At $t=0$, $x_1=0$, $x_2=0$, $v_1=2/3$, $v_2=1/3$, $\alpha=0$ and $\beta=1$. So the initial values of the second derivatives are $g_1=-5/18$ and $g_2=-2/9$. The initial slopes are $dv_1/dx_1=-5/12$ and $dv_2/dx_2=-2/3$.

From equation 20, v_2 may be solved as

$$v_2=(1/M)[R_1(1+x_1)x_1+(L_1+M)v_1-e_1] \qquad (24)$$

For the numerical example, equation 22 gives

$$g_2-2g_1=(0.5+x_1)v_1$$

Substituting the expression $g_2=2g_1+(0.5+x_1)v_1$ into equation 23 gives

$$g_1=\frac{0.5v_2-2\alpha\beta v_2^2+(\beta+1)(0.5+x_1)v_1}{-(2\beta+1)}$$

The v_1-x_1 and v_2-x_2 trajectories are shown in Fig. 5; the g_1-x_1 and g_2-x_2 trajectories in Fig. 6; and the instantaneous curves $x(t)$, $v(t)$ and $g(t)$ in Fig. 7.

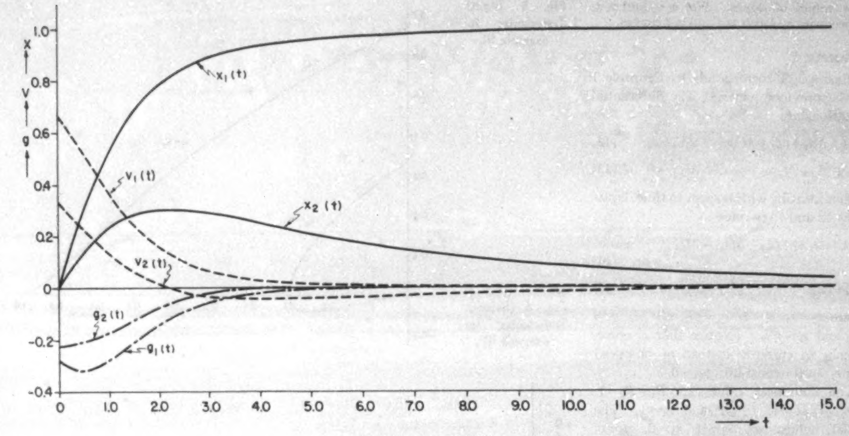


Fig. 7. Instantaneous curves, example III

EXAMPLE IV

The next example is given by considering a coupled circuit with a nonlinear capacitor in the secondary loop. The nonlinear capacitance C_N is assumed to have a constant part C_2 and a variable part Ke proportional to the voltage e across the capacitor. The differential equations, Fig. 8, are

$$(R_1+R_2)i_1+q_1/C_1-R_2i_2=e_1 \qquad (25)$$

$$-R_2i_1+R_2i_2+q_2/C_N=0 \qquad (26)$$

where

$$C_N=C_2+Ke \qquad (27)$$

Let $x_1=q_1$, $x_2=q_2$, $v_1=i_1=pq_1$ and $v_2=i_2=pq_2$, which gives

$$(R_1+R_2)v_1+x_1/C_1-R_2v_2=e_1 \qquad (28)$$

$$-R_2v_1+x_2/C_N+R_2v_2=0 \qquad (29)$$

Since $e=(i_1-i_2)R_2=(v_1-v_2)R_2$, where R_2 is the coupling resistance

$$C_N=C_2+KR_2(v_1-v_2) \qquad (27A)$$

Multiplying both sides of equation 29 by C_N and substituting in the expression given in equation 27(A) gives

$$-C_2R_2(v_1-v_2)-KR_2^2(v_1-v_2)^2+x_2=0 \qquad (30)$$

Putting in numerical values $R_1=R_2=1$, $C_1=C_2=1$, $K=1$, and $e_1=1$, from equations 28 and 30 gives

$$2v_1+x_1-v_2=1 \qquad (31)$$

$$-(v_1-v_2)-(v_1-v_2)^2+x_2=0 \qquad (32)$$

Differentiating with respect to time gives

$$2g_1+v_1-g_2=0 \qquad (33)$$

$$-(g_1-g_2)-2(v_1-v_2)(g_1-g_2)+v_2=0 \qquad (34)$$

Solving for g_1 and g_2 simultaneously gives

$$g_1=-\left[v_1+\frac{v_2}{1+2(v_1-v_2)}\right] \qquad (35)$$

$$g_2=-\left[v_1+\frac{2v_2}{1+2(v_1-v_2)}\right] \qquad (36)$$

Since $g_1=dv_1/dt$ and $g_2=dv_2/dt$, the following equation may be written for the v_2-v_1 or i_2-i_1 trajectory

$$\frac{dv_2}{dv_1}=\frac{g_2}{g_1}=\frac{v_1+2v_2+2v_1(v_1-v_2)}{v_1+v_2+2v_1(v_1-v_2)}=f(v_1,v_2) \qquad (37)$$

So, the result is an equation for the v_2-v_1 trajectory involving a function of v_1 and v_2 on the right-hand side of equation 37. In electrical terminology

$$\frac{di_2}{di_1}=\frac{i_1+2i_2+2i_1(i_1-i_2)}{i_1+i_2+2i_1(i_1-i_2)} \qquad (37A)$$

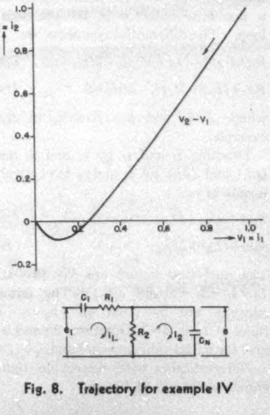


Fig. 8. Trajectory for example IV

The v_2-v_1 or i_2-i_1 trajectory is shown in Fig. 8.

The initial current values are $i_1=i_2=1$. The initial slope of the i_2-i_1 trajectory is 3/2 or 1.5.

Let dv_2/dv_1 be denoted by a symbol s, for slope. Then

$$v_2=\frac{(1-s)v_1(1+2v_1)}{(s-2)+2v_1(1-s)} \qquad (38)$$

Isoclines can be constructed from this relation. For $v_1=1$ and $s=1.5$, the result is $v_2=1$.

The initial values for $g_1=di_1/dt$ and $g_2=di_2/dt$ are $g_1=-2$, $g_2=-3$. According to equation 37 the slope is given by $g_2/g_1=1.5$.

Conclusion

An alternative method for the solution of nonlinear coupled circuits is here developed. This method is not only useful in checking the results given in the previous paper[1] but is also powerful in handling nonlinearities in both meshes of a 2-loop coupled circuit and, in general, for handling nonlinear elements in many meshes of a multiloop system.

Though simple examples in 2-loop coupled circuits are given for illustrations in this paper, it is believed that the method can be easily extended to the solution of two groups of nonlinear differential equations related by "coupling" terms between the two sets of variables in the two groups.

Take a simplified nuclear reactor problem for illustration. If six delayed neutron emitters are assumed, there is a group of six first-order differential equations coupled with the neutron density w, while another first-order differential equation involves w and is coupled with a group of variables $x_1, x_2, \ldots x_6$ belonging to the set of six differential equations. While the elimination of x's may result in a seventh-order nonlinear differential equation in w, the alternate method may prove simpler in actual calculations.

The method of solving simultaneous phase plane equations is, of course, closely related to the method of high-order phase planes. For the solution of two second-order differential equations in a coupled circuit with R, L, C in both loops, according to the method of high-order phase planes, a fourth order differential equation results. There are four derivatives involved besides the variable itself. It takes a 4-dimensional space to represent the variable and its first, second and third derivatives. In the method of simultaneous phase plane equations, there are two 2-dimensional planes to represent the v_1-x_1 plane and the v_2-x_2 plane. However, x_1, x_2, v_1, and v_2 are related in a 4-dimensional space. Instead of working with the 4-dimensional space, the relations between the two 2-dimensional planes are considered, giving two other 2-dimensional planes (also situated in space) relating v_1 and v_2, and also x_1-x_2. Thus there are for this case 4-phase planes v_1-x_1, v_2-x_2, v_2-v_1, and x_2-x_1. There are also then two sets of $v-x$ trajectories, (as also obtained in the solution by the previous method), one set of v_2-v_1 trajectory, and one set of x_2-x_1 trajectory. Thus in fact there are four first-order phase plane equations representing the relations between v_1, x_1, v_2, and x_2. In other words, in spite of the simple procedures outlined in this paper, high-order phase planes or phase space is implied in this method just as in the previous method.

References

1. Analysis of Nonlinear Coupled Circuits, Y. H. Ku. *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 626–31.

2. Acceleration Method Plane for Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance, Y. H. Ku. *Ibid.*, pp. 619–26.

3. A Method for Solving Third and Higher Order Nonlinear Differential Equations, Y. H. Ku. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 256, no. 3, Sept. 1953, pp. 229–44.

4. Analysis of Nonlinear Systems with More Than One Degree of Freedom by Means of Space Trajectories, Y. H. Ku. *Ibid.*, vol. 259, no. 2, Feb. 1955, pp. 115–31.

Minimizing and Mapping Sequential Circuits

W. S. BENNETT
NONMEMBER AIEE

SHANNON'S original paper suggesting a modification of the classical Boolean algebra of symbolic logic to a form of switching algebra led to the development among switching and computer designers of many special techniques.[1–3] Recent interest in the field is rapidly developing methods well within the reach of the occasional switching engineer who finds a moderately complex relay circuit necessary to the operation or sequencing of another design. Reference is made to these papers for a brief summary of the binary nomenclature and operators of switching algebra.[4,5] The approach suggested here assures the designer that he is manipulating an adequate number of relays, and provides a pictorial work sheet that keeps track of all contact positions. The work sheet permits easy recognition of possible holding circuit difficulties. Negation is proposed as the switching-algebra operator to denote shunt contacts, and is consistent with the mapping procedure.

The conventions used in this paper are:

Capital letters = switches or relay coils.
Lower case letters = contacts on corresponding relays.
a = normally open contact on A.
a' = Not a = normally closed contact on A.
1 = value of energized coil or closed contact, continuity.
0 = value of de-energized coil or open contact, no continuity. (If $A=1$, $a=1$ and $a'=0$; if $A=0$, $a=0$ and $a'=1$).
$ab'=a$ and b': contacts in series.
$a+b'=a$ or b': contacts in parallel.

Stating the Problem, Minimization

Minimizing the number of relays in a circuit may not lead to the simplest configuration but it is a common objective and generally correct. Consequently it should be the designer's first step. Once a circuit has been designed, the problem of going back and simplifying it, by elimination or combination, is made more difficult because the designer's thinking is bound to the original plan. Initial itemization of the inputs and required modes of operation will yield the minimum number of switching devices n needed. Modes M are defined as the various discrete situations, combinational or sequential, which the circuit must be able to distinguish.

Consider example 1: Two devices X and Y are operated by switches P and Q. Closing P then Q is to operate X, for the opposite sequence, Y. (The circuit is not required to recognize theoretical simultaneous closures of P and Q.) Five modes are presented as itemized sequentially in Table I.

Four combinational inputs are available from (P, Q): (0,0), (0,1), (1,0), and (1,1). These will be called external signals E. Since M is greater than E, a sequential circuit is necessary, incorporating memory. In the simple example here this is immediately apparent as the requirement to distinguish mode 3 from mode 5. To keep the initial specification of the problem as general as possible, let it be assumed that all operations are performed by relays, with the inputs E used to control the relays. If an identity

$$f(A)=p \qquad (1)$$

Table I. Discrete Modes of Example 1 as Determined by Switches P and Q

Mode	P	Q	Remarks
1	off	off	Rest mode
2	on	off	
3	on	on	*P* first
4	off	on	
5	on	on	*Q* first

Paper 55-436, recommeuded by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted October 19, 1954; made available for printing April 6, 1955.

W. S. Bennett is with the Operations Research Office, The Johns Hopkins University, Chevy Chase, Md.

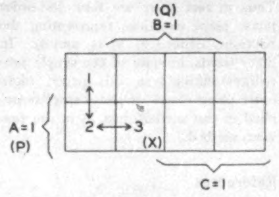


Fig. 1. First steps in sequential mapping of example 1. Arrows show reversible transitions from modes 1 to 2 and 2 to 3. Mode 3 (ABC′) represents the combination desired for operation of device X: P on, Q on, P first

for example, can be retained throughout the circuit development, then switch *P* can completely replace relay *A* (provided the switch can carry the required contacts) but such simplifications will not be incorporated until the end of the synthesis.

Since *n* relays can represent 2^n modes, five modes will require three relays. For example 1, then

$$n \geq 3 \qquad (2)$$

is a necessary condition. It must still be shown that it is sufficient. Each relay requires two signals for its operation: an "on" signal and an "off" signal (the "off" requirement assumes that the circuit is to be able to return to "rest"). (The methods presented here are applicable only to simple on-off direct-action relays and not to latching devices, steppers, etc., except as they may be represented by combinations of basic relays.) The three relays known to be necessary, then, for example 1, require six signals, of which only four are supplied by *P*, *Q*. Two additional internal signals, I, must be generated within the circuit to provide the desired sequential operation, i.e.

$$I = 2n - E = 2 \times 3 - 4 = 2 \qquad (3)$$

The net contribution of *n* relays is thus only $2^n - 2n$ states, because of the demands of the relays themselves. But this, plus the four external signals, is more than enough to satisfy the present 5-mode requirement

$$2^n - 2n + E = 2^3 - 2 \times 3 + 4 = 6 \qquad (4)$$

It is seen, then, that the number of relays both necessary and sufficient to distinguish *M* modes is found as the smallest value of *n* that satisfies

$$M \leq 2^n - 2n + E$$

$$\text{or } 2^n - 2n \geq M - E \qquad (5)$$

There will be many cases where the first estimate of *n*, based on *M* alone, will not be sufficient. For example, if $M = 7$, it might appear that three relays would suffice, providing eight modes. But if *E* were only 4, equation 5 would show that *n* must be 4. Inspection, and reference to equation 3, reveal that the requirement for internal signals in the memory circuits causes the increase in *n*. Note that no generalization has been lost: the value of *n* from equation 5 is sufficient even if a given relay is operated more than once, as will be seen in working out example 1.

Circuit Design by Sequential Mapping

Armed with the assurance that the minimum number of relays is known, the designer can utilize all available combinations in programming the actual sequencing. Continuing the discussion of example 1, the eight possible combinations of three relays are represented conveniently by a "map" or Venn diagram of the type used by Karnaugh.[6] The map, Fig. 1, can be used directly as a work sheet in programming the five required modes and two internal signals. The five original conditions may be called stable modes while the internal signals are generated by additional unstable modes. The example thus requires

$$M + I = 5 + 2 = 7 \qquad (6)$$

of the eight squares on the map. The eighth square (a "don't-care" condition) is available to the designer as an additional internal switching signal if wanted.

Fig. 1 has the lower four squares annotated $A = 1$, to denote that they represent the four possibilities in which coil *A* can be energized. The center four squares are similarly labeled $B = 1$, while the four on the right represent $C = 1$. Each of the eight possible combinations of the relay states *A*, *B*, and *C* is thus uniquely represented on the map. Since the type of replacement indicated by equation 1 seems a reasonable possibility, a *P* appears by *A* and a *Q* by *B* to remind the designer to try to permit this simplification during the mapping process. In other words, if *A* is operated directly by *P*, and only by *P*, then *A* may be unnecessary: all its functions can be performed directly by *P*, adding contacts to *P* if necessary. The rest mode is logically chosen as existing when all three relays are de-energized, i.e., $A'B'C'$. This square carries the figure 1 to identify it as mode 1; see Table I.

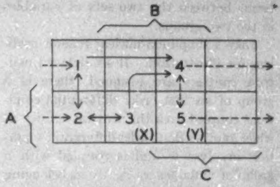


Fig. 2. Example 1 mapping completed. A′BC′ is an unstable mode which can be entered from mode 1 or mode 3 but in either case causes irreversible transfer to mode 4. Both B′C modes are used as unstable squares. Note that end squares are adjacent in the sense that direct passage can be made, e.g., from AB′C to AB′C′

The process of circuit design now consists of successively working through Table I (since it was originally set up to indicate one possible sequence of operation), recording the procedures and modes on the map; i.e., recording the route taken. Mode 2 was *P* on, *Q* off, which is seen to be the square $AB'C'$ since it is assumed that closing *P* operates *A*. A "2" is written in this square and the passage from mode 1 to mode 2 is noted by the arrow from 1 to 2. Since opening *P* should return the circuit to rest, the arrow is made double-ended and the transition will be called reversible. Referring to Table I for the next step in the sequence, it is seen that mode 3 calls for *Q* on, (*P* already on). Since Mode 2 had *P* already on, mode 3 is simply a transfer to the adjacent square ABC', and the "3" is recorded on the map. A reasonable decision which the designer must now make is that this transfer be reversible: the arrow from 2 to 3 is thus marked with a head at both ends. Mode 3 corresponds to the original requirements for the operation of device *X*: an *X* is accordingly marked in this square.

Note that the three squares thus far occupied all represent stable modes. Entry to or exit from such squares can

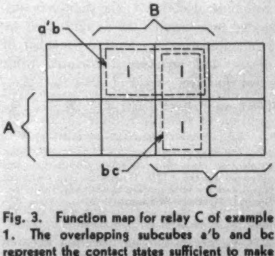


Fig. 3. Function map for relay C of example 1. The overlapping subcubes a′b and bc represent the contact states sufficient to make f(C)=1 in the designated squares

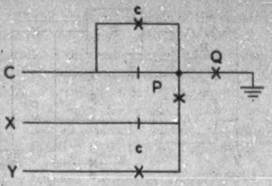


Fig. 4. Short schematic of solution to example 1: P is a transfer switch; Q is normally open. (—X— represents normally open contact; —|— normally closed contact). Relay C has two poles: one normally open, the other a transfer. All switching is shown in the ground returns; the supply voltages for relay C and devices X and Y are not shown. It is required only that they have a common ground

occur only as the result of external signals. In this sense they were stable conditions since the circuit would remain indefinitely in one of these conditions until a change occurred in the condition of P or Q.

The need for the memory contributed by relay C is seen as soon as the effort is made to map mode 4. Using mode 3 as a starting point (P on, Q on), consider the consequences of turning off P. If this is taken as passage to $A'BC'$ (and as passage to a stable state), then note that if this is considered reversible that the return passage to ABC' merely re enters mode 3 (operates X), whereas Q on should operate Y (mode 5). These difficulties are overcome by considering $A'BC'$ as an unstable mode, using it as an internal signal to energize relay C. ABC' is exited via $A'BC'$ which causes irrevocable entry into $A'BC$, designated "4" in Fig. 2.

Square $A'BC'$ therefore represents a transient state of the relays, i.e., an unstable mode. The circuit cannot remain in this condition, which includes $C=0$, because as soon as this exists it is to cause $C=1$. Obviously, the contact equation for the circuit of C when it is written must contain the condition

$$f_c(C)=a'bc' \tag{7}$$

The curved arrow of Fig. 2, from ABC' to ABC, summarizes all of the foregoing. Whenever the relay state $A'BC'$ exists, C will be given an energization path, causing entry into $A'BC$, so the passage through $A'BC'$ is not reversible: the arrow carries an arrowhead on only one end. $A'BC'$ can be exited only one way, but note that it can be entered from mode 1 as well as from mode 3. It is still unstable, but this additional path (the more obvious situation of Q on first, from rest) is represented by its own arrow. This arrow, too, is single-ended. Internal signals then may be generated in more than one manner, but will always cause passage to the same stable mode.

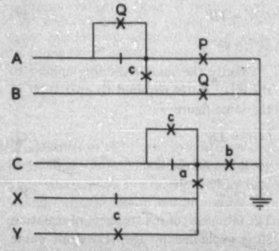


Fig. 5. Circuit of example 1 when 7-square solution is used instead of allowing don't-care square to be employed. Note added complexity over Fig. 4 which accomplished same functions

At this stage inspection of Fig. 2 shows that four of the five required stable modes have been filled and one of the two internal signals has been generated. The fifth mode can be entered as

$$f_r(Y)=ABC \tag{8}$$

but it is clear that both of the remaining squares are needed for convenient return to rest. One of these two (it is not necessary to try to say which one) was known to be needed, as the second internal signal; the other is now recognized as the expected don't-care square and consequently is available to the designer. The final map form can then be filled out as by the dashed arrows of Fig. 2. At no time was the operation of A and B compromised from the desired identification with P and Q. However, if the designer had not realized the use of the eighth square he would be forced to give up this simplification.

Equations: Transfer-Time Problems

The equations specifying the actual operating circuits of the relays can be written by inspection of the map, as suggested by Karnaugh.[5] With practice, these equations can be written at sight but the designer may wish to go through the intermediate step of sketching Karnaugh's function map. Fig. 3 represents such a sketch for relay C. Note that a "1" appears in the unstable square $A'BC'$, since this is the internal signal to energize C, equation 7, but none appears in either of the $B'C$ squares, since these

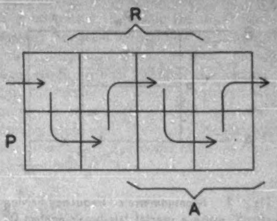


Fig. 6. Map of frequency-halving circuit, example 2. Stable modes are recognized as those containing arrowheads, and need not be labeled as was done in Figs. 1 and 2

are meant to be internal signals to cause the release of C. The relay states to operate C are

$$f_r(C)=A'BC'+A'BC+ABC \tag{9}$$

But $A'BC+A'BC=A'BC$ (10)

i.e., $1+1=1$

so $f_r(C)=A'B(C'+C)+BC(A'+A)$ (11)

or $f_r(C)=A'B+BC$ (12)

By the simple expedient of replacing capital letters by lower case letters, equation 12 is transformed to equation 13 which represents the contact states needed in series with C

$$f_c(C)=a'b+bc=b(a'+c) \tag{13}$$

Similarly $f_c(X)=abc'$ (14)

and $f_c(Y)=abc$ (15)

Fig. 4 represents the final circuit, using the short schematics of Keister, *et al.*[3] Note that the contacts associated with A and B from equations 13 through 15 have been labeled P and Q, since it is assumed the switches are able to carry sufficient contacts, making the relays unnecessary.

It is interesting to compare the circuitry necessary if the designer had not taken advantage of the don't-care square. Fig. 5 shows the 7-square solution, where $AB'C$ was not used, and is an impressive result of ignoring Karnaugh's dictum to make subcubes as large as possible, and keep them as few as possible.

For an illustration of transfer-time problems and techniques, consider the frequency-halving circuit discussed by Montgomerie and Raspanti.[4,6]

Example 2:

Mode 1—Relay R is to be operated when a pushbutton P is pressed.

Mode 2—R is to remain on when P is released.

Mode 3—The next time P is pressed R must de-energize.

Mode 4—Remain de-energized when P is released.

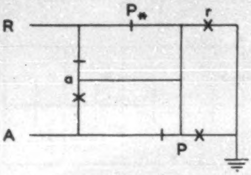


Fig. 7. Montgomerie's frequency-halving circuit, using a special late-break contact P* on the controlling switch P. Supply voltages for relays R and A not shown

These conditions establish four modes M and two external signals E, (P and P'). So

$$n=3 \tag{16}$$

$$I=4 \tag{17}$$

from equations 5 and 3. Of the three switching devices required, as established by the value of n, let P be one and R, itself, serve as another. Then only one auxiliary relay A is necessary and the circuit can be mapped as in Fig. 6.

Stable modes are excited only as the result of an external signal, here the operation or release of P, and occasion no difficulty during the contact transfer. However, unstable modes, generating internal signals, may permit erratic operation when the holding path of a relay is being transferred between contacts. Consistent use of overlapping subcubes in writing the equations frequently maintains circuit continuity during transfer time but questionable map areas should be checked systematically, as will be described. The relay-state equations of example 2, with overlapping subcubes, are written by inspection of Fig. 6

$$f_r(R)=A'P+RA'+RP' =(R+P)(A'+P') \tag{18}$$

$$f_r(A)=RP'+RA+AP =(R+P)(A+P') \tag{19}$$

Checking the four unstable squares of Fig. 6, consider $PR'A'$. This is to generate an internal signal to operate R, but no holding path is involved. Similarly, $P'R'A$ serves only to release A: no holding is necessary. But $P'RA'$ involves a holding circuit for R and PRA holds A, so these two cases will be evaluated.

Entry into $P'RA'$ results from the transfer P to P'. In evaluating $f(R)$ to see if the holding circuit is good, it is not safe to assume continuity through either p or p', since there will be a finite period of time when the transfer contact is in motion. The contact equation for R

$$f_c(R)=(r+p)(a'+p') \tag{20}$$

should be evaluated for the situation where p and p' are simultaneously zero

$$f_c(R)=ra' \tag{21}$$

Note that this term is due solely to the overlapping subcube term RA'. Equation 21 will have the desired value of unity only if both terms on the right have unit value. It is necessary to check only a'.

$$a'=1 \tag{22}$$

if

$$f_c(A)=(r+p)(a+p')=0 \tag{23}$$

under the conditions

$$p=p'=0 \tag{24}$$

Substituting equation 24 into 23

$$f_c(A)=ra \tag{25}$$

Equation 25 is consistent since it will have a value of zero if $a=0$; and $a=0$ if $F(A)=0$. Therefore, square $P'RA'$ presents no difficulty.

Next, evaluating PRA, the same condition, equation 24, must be considered in regard to the holding circuit of relay A. (Note that PRA does not hold R; in fact its function is to release R). Equation 25 contains the restrictions on p and p' and can have the value unity only if $r=1$. This requires

$$f_c(R)=ra'=1 \tag{26}$$

and is inconsistent, since a' cannot be unity in equation 26 and a be unity in equation 25 at the same time. In other words, A will release, not hold, in PRA; special techniques are necessary to prevent this occurrence. Montgomerie[6] proposed that a late-break contact be used on P to hold R through the starred contact in Fig. 7. This would retain the ar path for A until ap was established, when P^* could break without affecting A (R would release as required, though slightly delayed). Shunting the coil of R with a few microfarads' capacitance would accomplish the same thing, delaying the release of R momentarily but avoiding the need of special contacts.

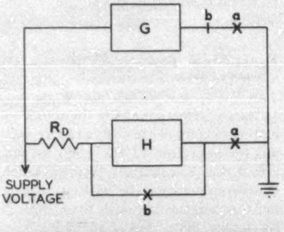


Fig. 8. The relay-state equations for G and H are identical, $f_r(G) = F_r(H) = AB'$ but although $f_c(G) = ab'$, the parallel network in the circuit for H is described by the partial dual: $f_c(H) = a(b')'$. The dropping resistor R_D serves to protect the supply voltage against a short circuit through ab

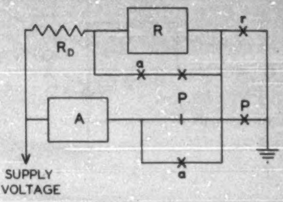


Fig. 9. Alternate solution of the frequency-halving problem of example 2, using shunt contacts instead of special late-break contacts, to avoid transfer-time difficulties. This circuit requires one more pole on A than Fig. 7 but has the same number of contact pairs

SHUNT CONTACTS

Raspanti[4] suggests shunting R with contacts. Series contacts cannot start to energize a coil until they are completely operated, i.e., at the end of the transfer period. However, normally closed contacts in parallel with a coil will permit the coil energization to start the instant they break: at the beginning of the transfer period. Such operation requires a dropping resistor, etc., but the technique is a powerful one and can be incorporated readily into the switching algebra. Negation, with a prime, is operationally and algebraically consistent and appropriate for shunt contacts. Its use is proposed as a substitute for Raspanti's complex algebra.

Consider the equations for relay G in Fig. 8: G is energized if and only if A is operated and B is not operated, viz.

$$f_r(G)=AB' \tag{27}$$

$$f_c(G)=ab' \tag{28}$$

Exactly the same reasoning applies to the relay states required to operate H in the same figure

$$f_r(H)=AB' \tag{29}$$

yet it is not true that

$$f_c(H)=ab' \tag{30}$$

The falseness, or inadequacy, of equation 30 is explained by the statement made earlier, that the contact function can be obtained from the relay function by changing to lower case letters, if the contact network is in *series* with the device operated.

At this point use is made of the simple duality of binary switching algebra: parallel is the dual of series, addition the dual of multiplication, normally open the dual of normally closed, 0 the dual of 1, etc. To make the contact of B from equation 29 a parallel contact, the contact function is written by negating the affected term

$$f_c(H)=a(b')' \tag{31}$$

Equation 31 is the proper representation of the circuit of H in Fig. 8. The presence of the primed group indicates that the dual of the contacts within the group represents a network to parallel H. The designer must therefore retain the identity of the group and not substitute the identity

$$(b')'=b \tag{32}$$

into equation 31 since this would yield a function as false as equation 30.

Applying this method to example 2, note equation 20 and convert the second term to a parallel network

$$f_c(R)=(r+p)(a'+p')' \tag{33}$$

Checking the holding circuit for A in PRA under this new condition but still subject to the p and p' restrictions of equation 24, it must be shown that $f_c(A)$ from equation 25 has unity value. This in turn requires r be nonzero, i.e.

$$f_c(R)=r(a')'=1 \tag{34}$$

Equations 25 and 34 are compatible, so shunt contacts have solved the problem in PRA. Because the circuit has been changed, $P'RA'$ should be rechecked. It can be shown in the manner just enumerated that this unstable square is still free of difficulty. The revised circuit (equivalent to Raspanti's), is shown by Fig. 9.

Remarks and Conclusions

The minimization rule presented here gives assurance that enough relays are being considered in conventional circuit design and that none are superfluous. Selector switches, steppers, latching relays, etc., are not considered, but it is hoped the techniques may be extended to these devices soon. Maps are limited in convenient usefulness to circuits with four or fewer relays, although Karnaugh proposes extensions, including a 6-relay 3-dimensional map. Even four relays offer 16 combinations which the designer may put on paper instead of juggling mentally. The advantage in having the don't-care squares apparent encourages their optimum utilization. Transfer-time problems require a final check of the system but the concept of unstable squares assures the consideration of all potential trouble spots while the binary duality of the algebra permits the use of shunt contact techniques.

References

1. A SYMBOLIC ANALYSIS OF RELAY AND SWITCHING CIRCUITS, **Claude E. Shannon.** *AIEE Transactions*, vol. 57, 1938, pp. 713–23.

2. THE DESIGN OF SWITCHING CIRCUITS (book), **William Keister, A. E. Ritchie, S. H. Washburn.** D. Van Nostrand Company, Inc., New York, N. Y., 1951.

3. SYNTHESIS OF ELECTRONIC COMPUTING AND CONTROL CIRCUITS (book), **Staff of the Harvard Computation Laboratory.** Harvard University Press, Cambridge, Mass., vol. 27, "Annals," 1951.

4. A COMPLEX ALGEBRA FOR RELAY CIRCUITS, **Matthew Raspanti.** *Electrical Engineering*, vol. 72, Nov. 1953, pp. 992–93.

5. THE MAP METHOD FOR SYNTHESIS OF COMBINATIONAL LOGIC CIRCUITS, **M. Karnaugh.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 593–99.

6. SKETCH OF AN ALGEBRA FOR RELAY AND CONTACTOR CIRCUITS, **G. A. Montgomerie.** *Journal*, Institution of Electrical Engineers, London, England, vol. 95, pt. II, 1948, pp. 355–64.

7. THE SYNTHESIS OF SEQUENTIAL SWITCHING CIRCUITS, **D. A. Hoffman.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 57, nos. 2–3, 1954, Mar., pp. 161–90, Apr., pp. 275–303.

A Sine Function Resistor

K. L. NIELSEN
NONMEMBER AIEE

E. H. ROLAND
NONMEMBER AIEE

Synopsis: This paper describes a sine function resistor of rather unusual design. A linear resistance card bent to circular shape about an eccentric frame of very simple construction is employed to obtain a resistance variation that is within very close limits of true sine values. This resistor should find application in certain computer circuits, particularly in those of the balanced-bridge type. The same theory is also applied to the function $(1-\sin\varphi)^i$ $(i=1/2, 1)$.

THE resistor to be described utilizes the familiar shunted circuit of Fig. 1. The resistance R_t is a quadratic function of slider displacement and is usually expressed in the form

$$R_t=\frac{R_t}{4}(1-x^2) \tag{1}$$

R_t is the total (unshort-circuited) resistance, and x is the displacement of the slider from the center position B, expressed as a fraction of its maximum possible travel to the right or left of B, Fig. 1. (x varies from -1 to $+1$.)

If the shunted resistor of Fig. 1 is bent around a cylindrical frame of radius r, as indicated schematically in Fig. 2, the resistance R_t may be said to be very roughly proportional to the sine of the angle φ between the slider arm and the chord AC. For this case, equation 1 may be written

$$R_t=\frac{R_t}{4}(1-x^2)\doteq\frac{R_t}{4}\sin\varphi \tag{2}$$

This may also be expressed as

$$R_t=R_1\left(1-\frac{R_1}{2R_2}\right)\doteq\frac{1}{2}R_2\sin\varphi \tag{3}$$

where R_1 and R_2 are as shown in Fig. 2. (For 0 degrees $<\varphi<$90 degrees, R_1 is measured from end C, as shown; for 90 degrees$<\varphi<$180 degrees, R_1 is measured from end A to arm.) From equation 3

$$\sin\varphi\doteq\frac{R_1}{R_2}\left(2-\frac{R_1}{R_2}\right) \tag{4}$$

Equations 3 and 4 are introduced here since the mathematical work outlined in the section following is based upon these expressions rather than the more frequently used form of equation 1. It will be noted that x is no longer a linear function of φ for the modified resistor to be described.

For the resistor configuration of Fig. 2, the approximation is indeed very crude since $(1-x^2)$ departs from $\sin\varphi$ by 20 per cent or more at some points. However, it has been found that a better fit to the sine curve can be obtained by displacing the pivotal point of the slider arm by an amount k from the center of the frame, as indicated schematically in Fig. 3. In fact, it will be shown that a value for the offset k can be found that will permit R_t

---

Paper 55-430, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted June 11, 1953; made available for printing April 1, 1955.

K. L. NIELSEN is with the United States Naval Ordnance Plant, Indianapolis, Ind., and E. H. ROLAND is with the General Motors Corporation, Indianapolis, Ind.

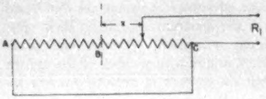


Fig. 1. Shunted resistor

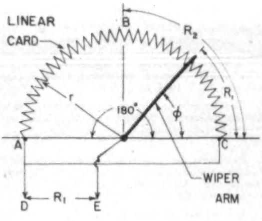


Fig. 2. Circular shunted resistor

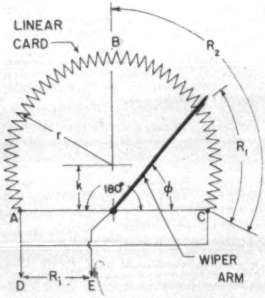


Fig. 3. Offset circular resistor

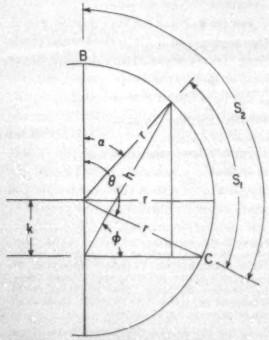


Fig. 4. Geometrical relations

Table I. Actual Values of θ

α°	0°	10°	20°	30°	40°	50°	60°	70°	80°	90°	100°	110°
θ°	all	114.025	113.980	113.938	113.896	113.872	113.862	113.895	113.982	114.146		

to be, within extremely close limits, proportional to sin φ. The problem encountered in the design of this resistor is that of obtaining the optimum value of the ratio of k to r, where r is the radius of the circular placement of the resistance card, for which R_1 is equal to $1/2R_2$ sin φ for every value of φ. This becomes a purely mathematical problem, the solution of which is discussed in the following section.

Derivation of Optimum Value of k/r

Since a linear resistance card is employed, the approximations can be written in terms of the arc lengths S_1 and S_2; see Fig. 4. Therefore the following may be written

$$\sin\varphi \doteq \frac{R_1}{R_2}\left(2-\frac{R_1}{R_2}\right) = \frac{S_1}{S_2}\left(2-\frac{S_1}{S_2}\right) \tag{5}$$

The following relations are easily obtained from Fig. 4

$$\sin\varphi = \frac{y+k}{h} \tag{6}$$

$$y = r\sin(90^\circ-\alpha) = r\cos\alpha \tag{7}$$

$$k = r\cos(180^\circ-\theta) = -r\cos\theta \tag{8}$$

$$\begin{aligned} h^2 &= r^2+k^2-2rk\cos(180^\circ-\alpha) \\ &= r^2+k^2+2rk\cos\alpha \\ &= r^2(1+\cos^2\theta-2\cos\theta\cos\alpha) \end{aligned} \tag{9}$$

The true value of the sine function of the angle φ is then given by

Table II. Check on $f(\alpha, 113.955^\circ) = g(\alpha, 113.955^\circ)$

α°	φ°	$\Delta(\alpha,\theta)$	Per Cent
0	90	0	0
10	82.8833	0.0000034	0
20	75.7399	0.0000096	0
30	68.5418	0.0000085	0
40	61.2586	−0.0000109	0
50	53.8558	−0.0000534	0.01
60	46.2929	−0.0001087	0.02
70	38.5214	−0.0001442	0.02
80	30.4815	−0.0001079	0.02
90	22.0981	0.0000456	0.01
100	13.2764	0.0002754	0.12
110	3.8962	0.0002600	0.38
113	0.9515	0.0000849	0.51
113.5	0.4542	0.0000425	0.54
113.6	0.3545	0.0000335	0.54
113.7	0.2548	0.0000243	0.55
113.8	0.1548	0.0000149	0.55
113.9	0.0550	0.0000053	0.56
113.955	0	0	0

$$\sin\varphi = \frac{\cos\alpha - \cos\theta}{\sqrt{1+\cos^2\theta-2\cos\theta\cos\alpha}} = f(\alpha,\theta) \tag{10}$$

From Fig. 4 also

$$S_1 = r(\theta-\alpha) \text{ and } S_2 = r\theta \tag{11}$$

The approximation therefore reduces to

$$\begin{aligned} \frac{S_1}{S_2}\left(2-\frac{S_1}{S_2}\right) &= \frac{r(\theta-\alpha)}{r\theta}\left(2-\frac{r(\theta-\alpha)}{r\theta}\right) \\ &= \left(1-\frac{\alpha}{\theta}\right)\left(2-1+\frac{\alpha}{\theta}\right) \\ &= 1-\left(\frac{\alpha}{\theta}\right)^2 \\ &= g(\alpha,\theta) \end{aligned} \tag{12}$$

From equation 8

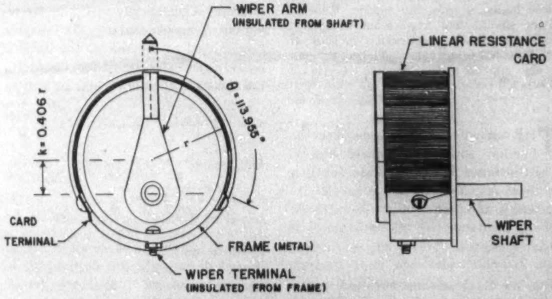


Fig. 5. Offset circular resistor construction

Vol.
Pt. 1
1955

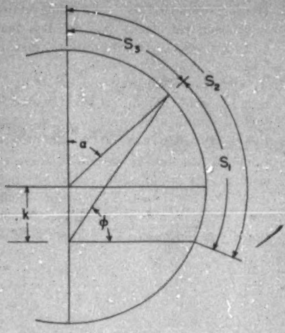


Fig. 6 (left). Geometrical relations

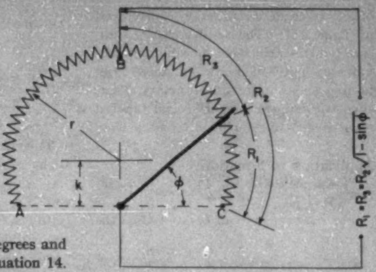


Fig. 7 (right). Square root resistor

$$\frac{k}{r} = -\cos\theta \qquad (13)$$

so that the problem of finding the optimum value of this ratio reduces to the finding of the optimum value of the angle θ.

It is desired to find the conditions under which

$$\Delta(\alpha,\theta) = g(\alpha,\theta) - f(\alpha,\theta) = 0 \qquad (14)$$

and, if possible, to find the optimum value of θ for which equation 14 will hold for all α.

From equations 10 and 12 it is easily seen that the best approach to the solution of equation 14 is a method of successive approximations.

The solution is obtained by letting α vary over the range of 0 to 110 degrees, taking the value at every 10 degrees and then solving for θ to satisfy equation 14. The saving feature of the problem is that the values of θ all cluster in the interval 113.8 degrees $<\theta<$114.2 degrees if the end points are disregarded. The actual values are shown in Table I.

The arithmetic average value of θ turns out to be

$$\theta = 113.955° \qquad (15)$$

The corresponding value of the ratio k/r is then, by equation 13

$$\frac{k}{r} = 0.406019 \qquad (16)$$

Calculations of Approximation Errors

If the values given by equations 15 and 16 are accepted, the actual deviation of the approximation given by equation 12 from the true value of the sine function given by equation 10 may be determined. If equation 15 is substituted into equations 10 and 12 respectively

$$f(\alpha,\theta) = \frac{\cos\alpha + 0.406019020}{\sqrt{1.164851445 + 0.812038040\cos\alpha}} \qquad (17)$$

and

$$g(\alpha,\theta) = 1 - 0.252800829\alpha^2 \qquad (18)$$

The deviations are given in Table II. The percentage value of the deviation is the per cent of the true value $f(\alpha,\theta)$.

The resistor is symmetrical about $\varphi = 90$ degrees so the errors for the angle from 90 to 180 degrees are symmetrically the same. The results show extremely good agreement between the approximation and the true value of the sine function. Even at the maximum percentage error the approximation gives the exact value of the sine function to five decimal places. They certainly justify the statement that the approximation is theoretically nearly exact. Furthermore, if the angle φ be limited to a smaller interval than 0 to 180 degrees, it is possible to find a still better value of θ and corresponding value of k/r for such an interval.

Mechanical Construction

The construction of the resistor presents no mechanical difficulties. The resistor may be of any radius r so long as the optimum value of k/r is maintained. The construction may be made similar to that suggested by Fig. 5. The resistor could be adapted to a 360-degree rotation by means of a double wiper arm, since points A and C of Fig. 3 are tied together electrically. Obviously, this method will also furnish a resistor for the cosine function.

Accuracy Test Results

The theory just presented was tested by actual measurements on a resistor so constructed, and the results indicate that they are well within the limits of resistor resolution at all points. A table of values

Table III. Test Values in Sine Function Resistor

$r_t = {}^1/_2R_2 \sin\varphi$			${}^1/_2R_2 = 287.66$ ohms		
φ, degrees	Measured r_t, ohms	Desired r_t, ohms	Δ	Per Cent error	Resolution, Per Cent sin φ
0	0.0096	0	0.0096		
1	5.240	5.020	0.220	4.4	25.0
2	11.884	10.039	1.845	18.5	12.5
4	20.894	20.067	0.827	4.1	6.2
6	31.085	30.069	1.016	3.4	4.2
8	41.063	40.034	1.029	2.6	3.1
10	50.864	49.952	0.912	1.8	2.5
20	99.132	98.385	0.747	0.76	1.2
30	144.86	143.83	1.03	0.72	0.75
40	186.44	184.90	1.54	0.83	0.52
50	220.55	220.36	0.19	0.09	0.37
60	249.73	249.12	0.61	0.24	0.25
70	270.49	270.31	0.18	0.07	0.16
80	283.24	283.29	−0.05	−0.02	0.08
90	287.50	287.66	−0.16	−0.06	0.001
100	283.43	283.29	0.14	0.05	0.08
110	270.61	270.31	0.30	0.11	0.16
120	249.29	249.12	0.17	0.07	0.25
130	220.04	220.36	−0.32	−0.15	0.37
140	185.41	184.90	0.51	0.28	0.52
150	144.21	143.83	0.38	0.26	0.75
160	98.129	98.385	−0.256	−0.26	1.2
170	50.020	49.952	0.068	0.14	2.5
172	40.215	40.034	0.181	0.45	3.1
174	30.319	30.069	0.250	0.83	4.2
176	21.407	20.067	1.340	6.7	6.2
178	11.088	10.039	1.049	10.5	12.5
179	4.679	5.020	−0.341	−6.8	25.0
180	0.0370	0	0.0370		

of a very crude construction is displayed in Table III. The results clearly indicate that the principle is feasible and provides a simple and practical method of obtaining an accurate sine function resistance without the use of special tapered resistor cards or complicated construction.

Application to Function $(1-\sin\varphi)^i$ $(i=1/2, 1)$

The same theory can be applied to the function

$$(1-\sin\varphi)^i, \left(i=\frac{1}{2}, 1\right) \quad (19)$$

If equation 5 is solved for S_1 and letting $S_3=S_2-S_1$ (see Fig. 6), it is easy to obtain

$$S_3 \doteq S_2\sqrt{1-\sin\varphi} \quad (20)$$

This function can be obtained by a slight modification to the construction of the sine resistor. The modification, shown in Fig. 7, consists of a tap at the center position B and disconnecting the shunt which tied together A and C. The errors committed by this construction have been calculated and are less than 0.05 per cent over the entire range of values.

Since the square root resistor of Fig. 7 is not a shunt circuit, it can be used in several circuit arrangements as a potential divider to obtain a voltage proportional to $\sqrt{1-\sin\varphi}$. A pair of these resistors and a suitable isolation stage could also be employed in a simple squaring circuit to obtain a voltage proportional to the function $(1-\sin\varphi)$.

Impulse Noise in Narrow-Band FM Receivers

S. P. LAPIN
ASSOCIATE MEMBER AIEE

J. J. SURAN
ASSOCIATE MEMBER AIEE

THE relatively narrow band frequency modulation (FM) communication receivers utilized in mobile communication are subject to interference caused by impulse-type signals received in the antenna. Most of these signals are from automobile ignition systems, from both the vehicle containing the radio and surrounding vehicles. Such impulses are also obtained from a variety of other electrical equipment. These impulses shock-excite the selective circuits in the receiver and produce a transient response from these selective circuits which interferes with the intelligence to be received from the desired information signals.

Disturbances in an FM receiving system may be attributed to two effects. One is the amplitude modulation caused by the interfering signal, and the other is the FM due to relative phase changes which occur between the wanted and unwanted signals. Amplitude modulation will be a factor only when the amplitude variations occur below the operating level of the limiters. In the ensuing discussion, the amplitude modulation caused by a disturbing signal will be considered negligible (i.e., ideal limiting action will be assumed), and only the FM effects will be considered.

The phase and magnitude relations between a transient disturbance and a carrier signal may be represented by a phasor diagram for any instant of time, such as illustrated in Fig. 1(A). T and C are the transient and carrier phasors respectively, and γ is the angle between them. The phasor triangle of Fig. 1(A) is uniquely determined by knowledge of the two sides T and C and the angle γ. The instantaneous frequency deviation is given by the time rate of change of phase of the resultant phasor R relative to the carrier C

$$\frac{\delta\phi}{\delta t}=f(t)=\frac{\delta u}{\delta|T|}\frac{d|T|}{dt}+\frac{\delta u}{\delta|C|}\frac{d|C|}{dt}+\frac{\delta u}{\delta\gamma}\frac{d\gamma}{dt} \quad (1)$$

where $\phi=u(|T|, |C|, \gamma)$ and $f(t)$ is the frequency deviation of the resultant phasor as a function of time. From equation 1 it is apparent that the instantaneous frequency depends upon the time rate of change of the transient and carrier amplitudes, as well as upon the rate of change of the phase between them.

Fig. 1(B) illustrates the variation of the resultant phase angle ϕ with time, when $|T|$, $|C|$, and $\delta u/\delta\gamma$ are constant, and for the two conditions $|T|<|C|$, and $|T|$ approximately equal to $|C|$. It can be seen that even when the total derivatives in equation 1 are finite, an infinite rate of change in ϕ is possible when the transient amplitude is equal to the carrier amplitude.[1]

The block diagram of a typical narrow-band FM receiver is illustrated in Fig. 2 Selectivity of the receiver is caused mainly by the filter. The dispersion,[2] or approximate spreading, of an impulse by a linear network is given by the well-known indeterminacy equation[3]

$$\Delta t\times\Delta f\approx 1 \quad (2)$$

where Δt is the dispersion in time, and Δf is the "effective" bandwidth of the network. Since the bandwidth of the FM receiver is almost entirely determined by the filter characteristics, the response of the receiver to an impulse will be governed, essentially, by the response of the filter. If the intermediate frequency section is composed of linear amplifiers followed by limiters, the impulse-response disturbance of the receiver will be the result of the phasor addition of the transient response and the signal, as indicated in Fig. 1(A). For any linear network,

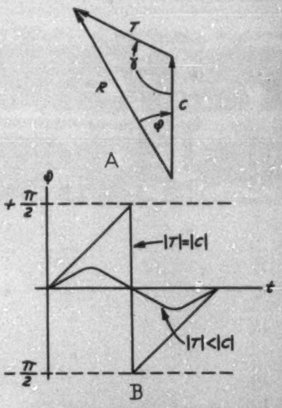


Fig. 1. A—Instantaneous transient-carrier phasor diagram. B—Variation of resultant phasor relative to carrier

---

Paper 55-429, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted July 12, 1954; made available for printing April 1, 1955.

S. P. LAPIN is with Motorola, Inc., Chicago, Ill., and J. J. SURAN is with the General Electric Company, Syracuse, N. Y.

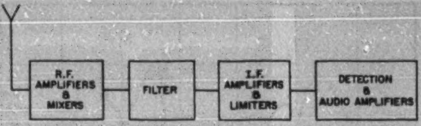


Fig. 2 (left). Block diagram of typical FM receiver

which is of minimum phase-shift character, the phase characteristic and transient response are uniquely determined by the attenuation characteristic. Thus, any combination of linear networks, which provides the same degree of selectivity as the filter in Fig. 2, will cause an equal disturbance at the discriminator when an impulse is applied at the antenna.

When two signals differ appreciably in amplitude, the signal of greater amplitude "takes over" in an FM receiver, and the receiver responds as if only the signal of greater amplitude were present. If the stronger interfering signal is constant in frequency, the receiver audio section will be blanked for the duration of this signal. When the amplitude of the disturbing signal approaches the amplitude of the information signal, the large frequency variations which may consequently result from the interaction of the two signals, caused by the sudden changes in phase, can cause shock excitation of the audio section. Under these conditions, the audible disturbance is due to the impulsive response of the audio section as well as to the frequency variations caused by the disturbance. Hence, when the interfering signal is either greater than, or approximately equal to, the amplitude of the desired signal, information will be lost. Therefore, the criterion for the design of the narrow-band filter in the receiver of Fig. 2, for the minimization of impulse-noise interference, must be to minimize the probability that the transient response of the filter will be of amplitude greater than or approximately equal to the amplitude of the desired signal.

The probability of generating large frequency excursions, during the critical interval of time when the disturbing signal is approximately equal to the desired signal, is proportional to the rate of change of phase between the transient signal and carrier signal, $d\gamma/dt$ in Fig. 1(A). This phase-impulse probability is also inversely proportional to the phase angle, γ_0, which is defined as the phase relation between the transient signal and the carrier signal at the instant of time the critical interval commences. For random impulse noise, γ_0 will vary randomly from one impulse to another. For example, if the rate of change of phase between the carrier and transient signals is 1,000 cycles per second, and the duration of the critical time interval is 25 μs per transient, the expected number of phase impulses per transient is 0.025. If the repetition rate of the transients is 100 per second, the expected number of phase impulses per second is 2.5.

For comparatively simple circuits, transient response may be calculated directly from the network transforms by means of operational techniques, (e.g., the Laplace transformation, etc.), but for more complex networks, such calculations become extremely involved. Hence, a simpler technique may be possible by assuming a transient response close to that obtained experimentally, and then finding the Fourier transform of this response. Generally

$$S(\omega)=F(\omega)F^*(\omega) \tag{3}$$

where $S(\omega)$ is the power spectrum and $F(\omega)$ and $F^*(\omega)$ are the Fourier inverse functions (direct and conjugate respectively) of the transient response $e_0(t)$.

Assumption of the transient response uniquely determines the network characteristic, $H(\omega)$, for the condition that the network is of linear, minimum phase-shift type. A mathematical function which was found to approximate closely the transients obtained experimentally from filters employed in typical narrow-band FM receivers is

$$e_0(t)=Kt^n\epsilon^{-\alpha t}\sin\omega_c t\,\frac{\sin^m(\omega_B t+\theta)}{(t-D)^m} \tag{4}$$

In equation 4, ω_c corresponds to the center frequency of the bandpass and is thus approximately equal to the frequency of the desired carrier in unichannel receivers, ω_b is related to the bandwidth of the channel ($BW\approx 2\omega_b$), the exponential argument α is a damping constant of the selective networks, D may be considered a delay term which is inversely related to the bandwidth ($D\approx 2\pi/BW$) and the other constants are determined by the number and effective "Q" of the particular selective circuits. An illustration of the wave form given by equation 1 is shown in Fig. 3. The Fourier integral of equation 4, multiplied by its complex conjugate, gives the circuit characteristic $S(w)$. The wave form of equation 4 is plotted in Figs. 3 and 4 for the values $n=1$ and $n=3$ respectively, where the following parameter values are assumed: $\alpha=20\times 10^3$ radians per second, $\omega_c=1.2\times 10^6$ radians per second, $\omega_b=30\times 10^3$ radians per second, $\theta=180$ degrees, $D=\pi/(30\times$

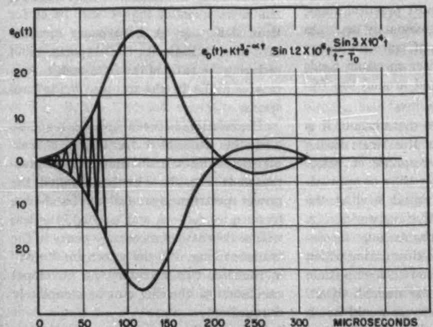


Fig. 3. Assumed transient response for N = 3

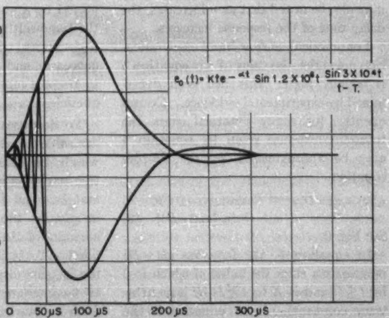


Fig. 4. Assumed transient response for N = 1

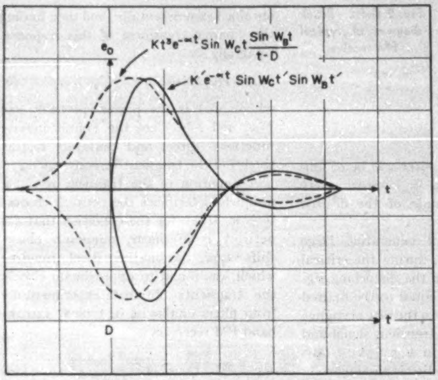


Fig. 5. Approximate and "exact" transient wave forms

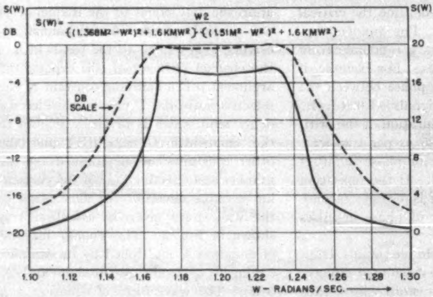


Fig. 6. Power spectrum associated with approximate transient response

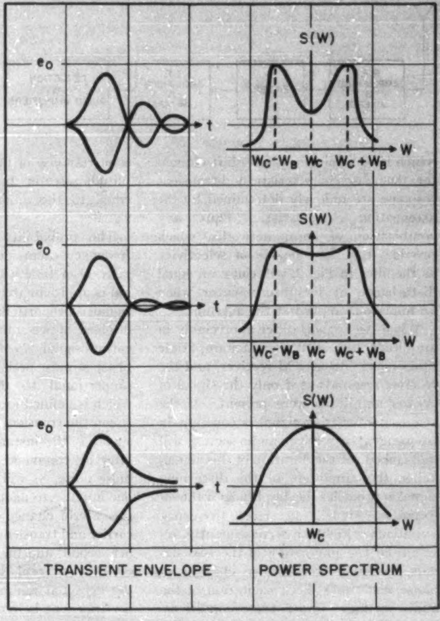


Fig. 7. Transient wave forms and associated power spectra

10^3) seconds and $m=1$. It is interesting to note that as n increases, the ratio of the second peak to the first peak in the transient response also increases. It may also be noted that as n increases, the delay time of the response increases.

For practical, narrow-band communication networks, the value of n in equation 4 is usually high. This has been determined by experimental evidence. Consequently, for many practical cases the transient response given by equation 4 may be approximated by the simpler equation

$$e_0'(t)=u(t-D)\epsilon^{-\alpha(t-D)}\sin[\omega_c(t-D)]\times \sin[\omega_B(t-D)] \quad (5)$$

See Fig. 5.

In equation 5, the function $u(t-D)$ represents a step, the value of which is 0 for $t \leq D$ and is K for $t \geq D$ (K is an arbitrary constant). From equation 5 the power spectrum associated with $e_0'(t)$ can be calculated rather simply; see the Appendix. The normalized spectrum associated with equation 5 is illustrated in Fig. 6. For any given bandwidth, the rate of decay of the transient is determined by α and, in most practical cases, the bandwidth is approximately $2\omega_b$. As α decreases, the rate of transient decay decreases, and the power spectrum tends to become more peaked around the frequencies $\omega_c-\omega_b$ and $\omega_c+\omega_b$.

To decrease impulse disturbance, it is desirable to decrease the time during which the transient response of selective networks is greater than or approximately equal to the signal level at the output of the selective networks. A measure of the disturbance may be defined on the basis of the time during which the impulse response is of such amplitude as to interfere with the desired signal. Such a disturbance measure would be on a per-impulse basis, and would serve as a means for comparing the relative disturbance effects of different transient wave forms. For example, suppose that the ratio of the maximum amplitude of the transient response to the signal level is 6 to 1. Experiment has indicated that the audio disturbance will continue until the noise amplitude is approximately one-half of the signal amplitude. Thus, the time, per noise impulse, that audio disturbance may be expected is 177.5μs for the wave form of Fig. 4 and 195μs for that of Fig. 3. A disturbance measure of this sort indicates the importance of reducing the ratio of the first peak to successive peaks in the impulse "ring" response.

The relation between impulse response and the associated frequency characteristic of the selective networks is illustrated in Fig. 7. The less peaked the power spectrum is around the band-edge frequencies ($\omega_c-\omega_b$ and $\omega_c+\omega_b$), the less will be the ratio of successive peaks in the transient ring. If the spectrum "nose" is rounded, the secondary (or envelope) oscillation of the ring can be completely damped out.

It is possible to photograph the response of the audio circuit in an FM re-

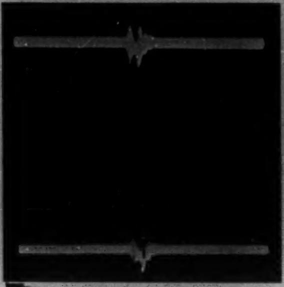


Fig. 8. Audio response due to transient of Fig. 10

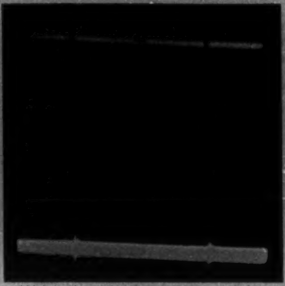


Fig. 9. Audio response for phase suppression circuit

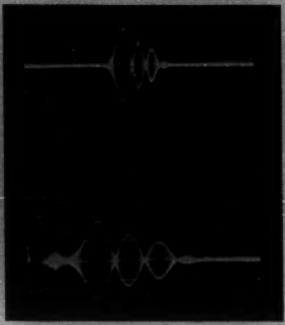


Fig. 10. Transient response of filter 10 kc wide

ceiver when impulses are applied at the antenna in the presence of a carrier signal. By varying the time exposures, audio response patterns showing the impulse excitation of the audio circuit by phase impulses and the blanking of the audio circuit by the high-amplitude constant-frequency portion of the transients become apparent. Photographs taken in the audio section of a Motorola unichannel receiver are shown in Figs. 8 and 9. Fig. 10 shows the impulse response of a filter 10 kc wide, photographed at sweep periods of 850μs and 250μs respectively.

Fig. 8 is a photograph of the audio response when this filter is used in the receiver. Since the rise time of the transient is longest at the leading end, the probability of phase impulses is a maximum in this region. This is verified by the associated audio response which indicates a maximum light intensity corresponding to the region of maximum phase-impulse probability. Both exposures in Fig. 8 were taken at sweep periods of 2,000μs; at the top is a 15-second exposure and at the bottom a 3-second one.

Fig. 9 illustrates the audio response of a receiver whose filter is contained in a phase-suppression feedback circuit.[4] Here the probability of phase impulses is theoretically reduced to zero. The top of Fig. 9 is a 3-second exposure which clearly indicates the blanking effect of the transient and the absence of shock-exciting phase impulses. The bottom of Fig. 9 is a 20-second exposure of the same circuit. Both were taken with a sweep period of 5,000μs.

The generation of phase impulses which cause shock excitation of the audio circuits in an FM receiver is manifested audibly by clicking sounds. Blanking of the desired signal by a disturbance of appreciably greater amplitude and constant frequency is heard as a popping sound. The two phenomena are generally referred to in the literature as "clicks" and "pops." Experimental evaluation tests indicate that most people find clicks more audibly disturbing than pops, and thus a receiver with a higher click probability may sound more annoying than another with a lower click rate. Furthermore, people seem to find high-pitched noises more disturbing than low-pitched ones, and thus clicks being equal, a receiver having a shorter pop duration may sound more annoying than one with a longer pop interval. However, the major consideration should be how the impulse response of a receiver interferes with a desired information signal and this, in general, is not the same as the manner in which a receiver reacts to noise alone.

Increasing the bandwidth of a receiver network will decrease the ring time of the transient response, and will therefore decrease the information disturbance for a given impulse repetition rate and for a given signal-to-noise power ratio. When the time duration of the transient response of the linear networks preceding the detection stage becomes equal to or less than the time duration of the transient response of the detection stage and succeeding audio networks, no further advantage is gained by increasing the intermediate frequency bandwidth.

Compensation of receiver networks to reduce impulse noise may result in an increase of random background noise, and vice versa. Thus, a compromise must generally be made in order to obtain an optimum operating point in relation to noise in general rather than relative to a particular type of noise. For example, as is evident from Shannon's equation[5] for the information capacity of a channel, if random noise is directly proportional to the bandwidth of a channel, the disturbance of information due to random noise will vary from a small amount to a finite maximum with increasing bandwidth. On the other hand, the disturbance due to impulse noise will vary inversely with increasing bandwidth. Thus, an optimum bandwidth must be defined which represents the best compromise between impulse-noise and random-noise disturbance. In addition, of course, the bandwidth specification must be such as to preclude the possibility of interference from other channels, as well as to satisfy the information capacity required by the desired signal.

Field tests were made with a receiver in the 152- to 174-megacycle 2-way-communication band. The vehicle was arranged so that rapid switching could be made between a standard receiver and one containing a filter with a frequency characteristic designed to minimize its transient response. In areas of relatively low signal strength, a definite improvement was noticed in the modified receiver as compared to the standard model.

Conclusions

Interference in a frequency-modulated receiver is manifested in two ways in high impulse-noise-to-signal areas. The first effect is blanking of the desired signal by a high amplitude, constant-frequency transient which results from the shock excitation of the predetection networks. Audibly, this blanking sounds like a pop. The second effect is due to rapid phase variation between the transient and desired signals, when they are nearly of the same amplitude, which gives rise to large and rapid frequency excursions at the discriminator. These frequency excursions result in shock excitation of the audio circuits, and are audibly apparent as clicking sounds. In relatively narrow-band receivers, the blanking, or pop effect, accounts for more loss of information than the audio shock-excitation, or click effect, although this latter effect may be more psychologically annoying. Time duration of the pops may be reduced by increasing the bandwidth or shaping the power attenuation characteristic of the selective networks in such a way as to reduce the selectivity of the channel.

Appendix I. Calculation of $e_0(t)$

Consider the Fourier integral

$$\int_0^\infty f(\theta)\epsilon^{-j\omega\theta}d\theta=F_a(j\omega) \tag{6}$$

Now, let $\theta=t-D$. Then

$$\int_D^\infty f(t-D)\epsilon^{-j\omega(t-D)}dt=F_a(j\omega) \tag{7}$$

or

$$\int_D^\infty f(t-D)\epsilon^{-j\omega t}dt=\epsilon^{-j\omega D}F_a(j\omega) \tag{8}$$

If the function $f(t-D)$ is zero from $t=0$ to $t=D$, equation 8 becomes

$$\int_0^\infty f(t-D)\epsilon^{-j\omega t}dt=\epsilon^{-j\omega D}F_a(j\omega) \tag{9}$$

Equation 9 is a consequence of the theorem of real translation.[6]

From this result, it is evident that the Fourier integral of equation 5 is

$$F(j\omega)=\epsilon^{-j\omega D}\int_0^\infty \epsilon^{-\alpha t}\sin(\omega_c t)\times \sin(\omega_B t)\epsilon^{-j\omega t}dt \tag{10}$$

In equation 10, ω_c represents a frequency which is equal to or approximately equal to the carrier frequency and it can be written as an integral multiple of ω_b. If

$$\omega_c=n\omega_B$$

where n is an integer, then

$$F(j\omega)=\epsilon^{-j\omega D}\int_0^\infty \epsilon^{-(\alpha+j\omega)t}\times \sin(n\omega_B t)\sin(\omega_B t)dt$$

$$F(j\omega)=\frac{2n\omega_B^2(\alpha+j\omega)\epsilon^{-j\omega D}}{[(\alpha+j\omega)^2+(n-1)^2\omega_B^2]\times[(\alpha+j\omega)^2+(n+1)^2\omega_B^2]}$$

$$F(j\omega)=\frac{2n\omega_B^2(\alpha+j\omega)\epsilon^{-j\omega D}}{[\alpha^2+(n-1)^2\omega_B^2-\omega^2+j2\alpha\omega]\times[\alpha^2+(n+1)^2\omega_B^2-\omega^2+j2\alpha\omega]}$$

The power spectrum is given by

$$S(\omega)=F(j\omega)F^*(j\omega) \tag{11}$$

and

$$F^*(j\omega)=\frac{2n\omega_B^2(\alpha-j\omega)\epsilon^{+j\omega D}}{[\alpha^2+(n-1)^2\omega_B^2-\omega^2-j2\alpha\omega]\times[\alpha^2+(n+1)^2\omega_B^2-\omega^2-j2\alpha\omega]}$$

Performing the multiplication indicated by equation 11 gives

$$S(\omega)=\frac{4n^2\omega_B^4(\alpha^2+\omega^2)}{\{[\alpha^2+(n-1)^2\omega_B^2-\omega^2]^2+4\alpha^2\omega^2\}\times\{[\alpha^2+(n+1)^2\omega_B^2-\omega^2]^2+4\alpha^2\omega^2\}} \tag{12}$$

For the following parameter values

$\alpha=20\times10^3$ radians per second
$\omega_c=1.2\times10^6$ radians per second
$\omega_b=30\times10^3$ radians per second

$$n=\frac{\omega_c}{\omega_b}=40$$

Equation 12 becomes

$$S(\omega)\approx\frac{5.18\times10^{21}(4\times10^8+\omega^2)}{\{[1.37\times10^{12}-\omega^2]^2+1.6\times10^9\omega^2\}\{[1.51\times10^{12}-\omega^2]^2+1.6\times10^9\omega^2\}} \tag{13}$$

Equation 13 is therefore the power spectrum of the transient-response wave form given by

$$e_0'(t)=u(t-D)\epsilon^{-20\times10^3(t-D)}\sin[30\times10^3(t-D)]\sin[1.2\times10^6(t-D)] \tag{14}$$

References

1. **Frequency Modulation Noise Characteristics**, **M. G. Crosby.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 25, Apr. 1937, pp. 472 ff.

2. **Theory of Communication**, **D. Gabor.** *Journal*, Institution of Electrical Engineers, London, England, vol. 93, pt. III, no. 26, Nov. 1946.

3. **Frequency Analysis, Modulation and Noise** (book), **S. Goldman.** McGraw-Hill Book Company, New York, N. Y., first ed., 1948.

4. **Reduction of Interference in FM Receiver by Feedback Across the Limiter**, **R. M. Wilmotte.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 40, Jan. 1952, pp. 34–36.

5. **Mathematical Theory of Communication** (book), **C. Shannon.** University of Illinois Press, Urbana, Ill., 1949.

6. **Transients in Linear Systems** (book), **M. F. Gardiner, J. L. Barnes.** John Wiley & Company, Inc., New York, N. Y., 1942.

Vol. 7
Pt. 1
1955

The Movable Core Transformer as a Pickup for Industrial Telemetering Systems

A. J. HORNFECK
MEMBER AIEE

L. M. WERMELSKIRCHEN
NONMEMBER AIEE

THE differential transformer is a transducer type of pickup in which mechanical displacement is converted into a proportional electric output by varying the mutual inductance between a primary winding and two secondary windings. The effect is a differential one so that motion of the core or rotor in one direction produces an increase in induced voltage in one secondary winding and a decrease in the other. A reversal of the direction of motion produces a reversal of the direction of change of secondary voltages. There are three basic types of differential transformers:

1. The mutual inductor type in which the flux linkage between the primary and secondary windings is varied by angular or linear displacement of the primary winding relative to the secondaries.

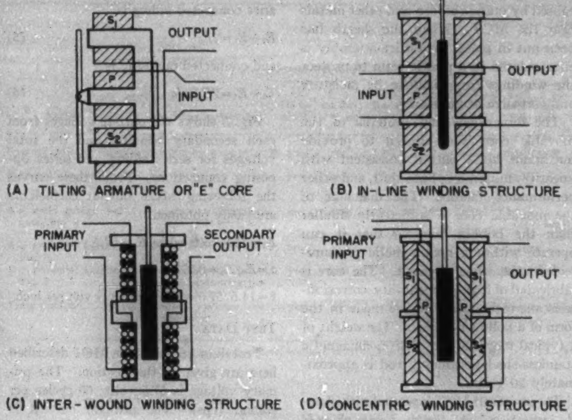
2. The type in which the coupling between the primary and two secondaries is differentially varied by linear or angular positioning of a short-circuited winding.

3. The movable core transformer (MCT) in which the flux linkages between the primary and each of the two secondaries is established by the position of the magnetic core. In the most common form, this device consists of a solenoid-type coil and a solid or laminated form of simple bar core mechanically linked to the primary measuring element.

The MCT is perhaps the most important of these types of differential transformer pickups for the telemetering and remote control of process quantities which are commonly measured by displacement-type elements such as floats, bellows, bourdon tubes, and diaphragms. Typical measurands are pressure, flow, level, and weight. Although there are earlier references in technical literature and patents, it appears that the first widespread application as a telemetering pickup of the MCT was made about 1940. The application of such devices to industrial process measurements was described by Dickey and Hornfeck.[1] Its first extensive usage coincides with the introduction of the electronically operated servo-balanced recorders to the measurement and control of industrial processes.[2]

Paper 55-438, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted November 24, 1954; made available for printing April 7, 1955.

A. J. Hornfeck and L. M. Wermelskirchen are with the Bailey Meter Company, Cleveland, Ohio.



(A) TILTING ARMATURE OR "E" CORE (B) IN-LINE WINDING STRUCTURE (C) INTER-WOUND WINDING STRUCTURE (D) CONCENTRIC WINDING STRUCTURE

Fig. 1. Movable core transformers

ol. 74
t. 1
955

The MCT is particularly adaptable to the null-balance circuit type of receiver instruments which permit wide range and high-accuracy measurements, independent of connecting line characteristics and loading on the instrument mechanism. During the past decade several instrument manufacturers have adopted the MCT for the electric telemetering of process variables represented by displacement functions.

Fig. 1 shows several basic forms of the MCT. Fig. 1(A) shows one form of the tilting armature type which has been used to some extent as a pickup element for very small motions in control applications. It is perhaps not suitable for accurate telemetering of measurements because of the extremely narrow range of motion for linear output and the magnetic forces on the armature which are difficult to balance out. Figs. 1(B), (C), and (D) illustrate the basic forms of the linear motion MCT which are now being manufactured by several companies.

The structure shown in Fig. 1(B) has been disclosed in several early patents[3,4] and is described by Schaevitz.[5] Fig. 1(C) is the form developed by MacGeorge[6] and claims to have several advantages in close magnetic coupling and accurate reproducibility. Its major disadvantage appears to be the rather large capacity coupling between primary and secondary windings. The concentric winding transformer shown in Fig. 1(D) is disclosed in several technical papers[1,7,8] and system patents.[9–11] It is the type preferred by the writers since its structure and characteristics make it suitable for the requirements of a simple and reliable pickup having a large signal output.

Features of the MCT Pickup

In spite of its structural simplicity the MCT is extremely versatile in its application and in the variety of circuits in which it can be used. The characteristics which make it uniquely suitable for industrial telemetering of position variables are summarized as follows:

1. The MCT has high signal output compared with other types of transducers such as strain gauges. For example, the output of the design described later in the paper is in the order of 1 to 10 volts as compared with approximately 10 millivolts for resistance strain gauges. In terms of power, the ratio between the output of the MCT and typical strain gauge pickups is in the order of 10^4 to 10^5.

2. No sliding contacts such as are required for slide-wire or selsyn devices are used. The moving core can be made free-floating which eliminates all frictional wear resulting in a device of nearly unlimited life.

3. At power frequencies up to at least 400 cycles, the moving core can be pressure-sealed in a nonmagnetic metal tube. This is an important feature in flow transmitters since the fluid may be at a high static pressure.

4. The output impedance is relatively low being of the order of a few hundred ohms which permits the use of simple wire-transmitting circuits without line-matching amplifiers at the sending end.

5. Linearity can be made to within 0.10 per cent over limited stroke ranges. Reproducibility between units is in the order of ±0.5 per cent.

6. Magnetic reaction forces are small and can usually be made negligible in terms of instrument accuracy.

7. One design can be used for a wide range of core motions. The MCT is suitable for full-scale motions from 0.025 to 0.800 inch.

8. The accuracy of the MCT when used in a null-type, servo-balanced system is very high, being a function of the geometric relation between core and coil, and is independent of supply voltage and frequency.

9. The isolation of the signal windings from the power supply permits the incorporation of a number of pickups in computing circuits without the necessity of separate power supplies for each element.

Design Data

The assembly of the MCT designed to satisfy the requirements for an industrial

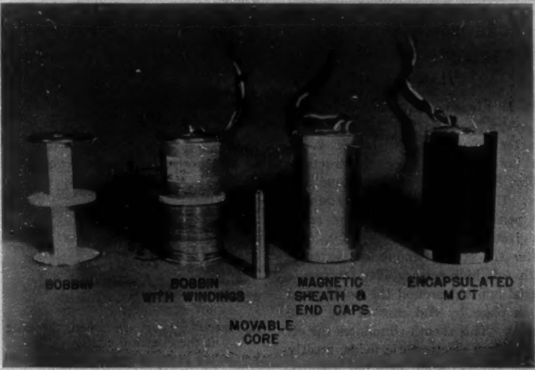
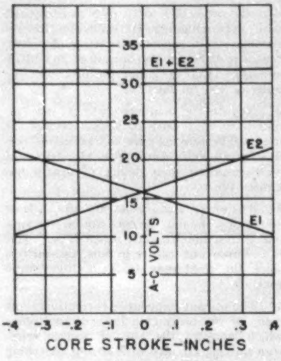
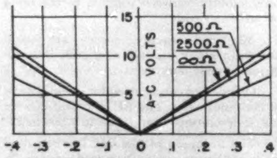


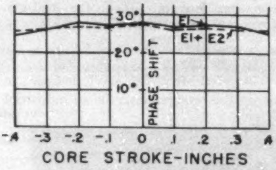
Fig. 2. Shielded and encapsulated MCT with rigid metallic bobbin



(A) INDIVIDUAL SECONDARY VOLTAGES



(B) TOTAL OUTPUT VOLTAGE AS A FUNCTION OF LOAD RESISTANCE, SECONDARIES CONNECTED SERIES OPPOSING



(C) SECONDARY VOLTAGE PHASE SHIFT

Fig. 3. Output characteristics of MCT

telemeter pickup is shown in Fig. 2. A metal bobbin is used to obtain structural rigidity, dimensional accuracy, and stability. Use of high-strength stainless steel permits maximum winding space for a given size transformer.

After the coil is wound, an external sheath of high-permeability magnetic material is fitted around the bobbin and held in place by end caps of the same material. This sheath eliminates the effects of stray electrostatic fields, greatly reduces effects from external electromagnetic fields, and eliminates distortion in the flux path of the MCT that might be caused by masses of iron and other metals near the MCT. When the sheath has been put in place, the entire assembly is encapsulated in a plastic resin to protect the windings from damage by moisture and corrosive atmospheres.

The dimensions and material of the movable core were chosen to provide maximum MCT output consistent with linearity, minimum phase shift, and other performance factors. The diameter of the movable core is sufficiently smaller than the bobbin bore so that it can operate within a nonmagnetic pressure-sealed tube when required. The core is fabricated of high-permeability, corrosion-resistant nickel alloy and is made in the form of a hollow cylinder. The weight of a typical movable core with nonmagnetic stainless-steel positioning rod is approximately 20 grams.

The standard MCT is wound with class-*A* insulation. A design using class-*H* insulation is also available and will exactly duplicate the performance of the lower temperature unit. Several variations of the basic coil structure have been made available. One of these, referred to as a double-function MCT, has two separate sets of secondary windings enabling the use of one transmitter in two independent circuits.

Performance Characteristics

The individual secondary voltages E_1 and E_2 can be expressed by the following equations for open-circuit conditions over the range of core travel in which the output of the MCT is linear

$$E_1 = \omega_0 M_0 E_s / Z_p (1 - m/M_0 x) \qquad (1)$$

$$E_2 = \omega_0 M_0 E_s / Z_p (1 + m/M_0 X) \qquad (2)$$

where

$\omega_0 = 2\pi f_0$ where f_0 is primary supply frequency
M_0 = mutual inductance (primary to secondary) when the core is at the electrical center
Z_p = primary impedance
m = change in mutual inductance per unit of core motion
E_s = primary supply voltage
x = core motion from the center of MCT

This may be expressed in terms of voltages

$$E_1 = aE_s(1 - k/ax) = E_0(1 - k/ax) \qquad (3)$$

$$E_2 = aE_s(1 + k/ax) = E_0(1 + k/ax) \qquad (4)$$

where

E_0 = secondary voltage when the core is at the electrical center
k = change in secondary voltage per unit of core motion
$a = \omega_0 M_0 / Z_p = E_0/E_s$

The total output voltage with secondaries connected aiding is

$$E_2 + E_1 = 2E_0 \qquad (5)$$

and connected opposing is

$$E_2 - E_1 = 2E_0 kx/a \qquad (6)$$

Fig. 3 shows the output voltages from each secondary winding and the total voltages for series aiding and series opposing connections. From these curves the following MCT output coefficients are easily obtained

$E_0 = 15.8$ volts when $E_s = 57.5$ volts

$a = E_0/E_s = 0.275$

$k = 14.6/57.5 = 0.254$ volts per volt per inch

Test Data

Test data taken for the MCT described here are given in this section. The primary voltage is 57.5 volts, 60 cycles per second (cps). The primary current is 0.13 ampere.* The primary watts are 4.4 watts* (7.5 volt-amperes).* The asterisk indicates that the core is positioned at the center of the MCT and that the secondaries are open-circuited.

Fig. 3(A) and (B) show the individual secondary output voltages E_1 and E_2 as well as total MCT output voltages with secondaries connected series aiding and series opposing. The secondary voltage phase shift is shown in Fig. 3(C) with respect to the supply voltage. Phase displacement between the two secondary voltages varies from 0 to 1 degree for a core stroke of 0 to 0.250 inch.

The MCT resistance and inductance is as follows:

Winding	D-C Resistance, Ohms	Inductance,* Henrys
Primary	150	1.0
Each secondary	105	0.3

The primary and secondary impedances are:*

1. Primary equivalent impedance:*

$Z_p = V_p/I_p = 440\underline{/70 \text{ degrees}}$ ohms

where

V_p = primary voltage
I_p = primary current

2. Each secondary equivalent impedance is:*

$Z_s = V_{0c}/I_{sc} = 150\underline{/45.5 \text{ degrees}}$ ohms

where

V_{0c} = open-circuit secondary voltage
I_{sc} = short-circuit secondary current

3. Secondaries connected series aiding are:*

Table I. Telemeter Systems Using the MCT Transmitter

System	Receiving Instrument	Power Required		Transmission Lines					Auxiliary Devices Required at MCT Transmitter
		At Transmitter	At Receiver	Wires Required	Maximum Resistance per Lead, Ohms	Shielding Required	Maximum Distance	Limiting Factor	
Remote indicator, (a-c)	a-c rectifier voltmeter	yes	no	2	50	yes	1,500 feet	line leakage and capacity	constant voltage supply for MCT primary
Remote indicator, (d-c)	d-c voltmeter	yes	no	2	50	no	1 mile	line resistance and leakage	regulated voltage supply for MCT primary rectifier, and filter on MCT output
A-c ratio bridge balance, Fig. 4	a-c servo with resistance slide-wire follow-up	no	yes	5	15	yes	determined by wire size	line resistance	resistor network for MCT core strokes exceeding 0.250 inch
A-c potential balance, Fig. 6	a-c servo with follow-up	no	yes	4	100	yes	3 miles	line leakage and capacity	resistor network for MCT core strokes exceeding 0.250 inch
D-c potential balance, Fig. 9	d-c servo with resistance slide-wire potentiometer	yes	yes	2	see note	no	50 miles	line resistance and leakage	a-c to d-c signal converter unit
Pulse duration	readout printer or indicator	yes	yes	wire or radio link maximum distance determined by transmission characteristics					scanner, coding and other transmission apparatus depending upon type of link

Note: The ratio of total resistance per lead to total leakage resistance must not exceed 0.01.

$Z_s{}' = V_{0c}/I_{sc} = 300\underline{/45.5 \text{ degrees}}$ ohms

4. Secondaries connected series opposing (with the core 0.250 inch from MCT center) are:

$Z_s{}'' = V_{0c}/I_{sc} = 230\underline{/24 \text{ degrees}}$ ohms

The MCT described is capable of delivering up to 0.75 watt into a 500-ohm load with the windings connected series aiding. Load impedances down to 5,000 ohms have negligible effect on the open-circuit calibration for either series aiding or series opposing secondary connections.

The primary winding temperature rise is 40 degrees Fahrenheit with the core in place. Excessive temperature rise will occur if the MCT is operated without the movable core.

External iron masses have negligible (less than 0.1 per cent) effect on the MCT calibration. An external a-c field of 60 cps with a strength of 300 ampere-turns will result in a 1-per-cent error if the field is in a direction to aid the primary ampere-turns. The error is less than 0.25 per cent if the external flux opposes the primary flux.

The reaction force on the moving core is zero when the core is centered but increases nonlinearly to a maximum of approximately 1 gram for a ±0.250-inch displacement, and is a function of the square of the MCT primary voltage.

Telemeter Circuits

A wide variety of telemetering circuits and receivers of the indicating, recording, or integrating type can be used with the MCT pickup. The selection of circuit and receiver depends on factors such as distance, type of transmission link available, and the accuracy and sensitivity of measurement required. Systems used may be classified as follows:

1. Remote-indicator system of direct-deflection type.
2. A-c bridge and potentiometer systems using automatic null-balance receiver instruments.
3. D-c bridge and potentiometer systems using null-balance receiver instruments.
4. Pulse-duration system for long-distance telemetering.

The first three systems are usable only with wire links. The fourth system can be used with either wire or radio links and can be multiplexed. A summary of the receivers and transmission-line requirements for these telemeter systems is given in Table I.

REMOTE INDICATOR

The simplest and perhaps most primitive system uses a receiver indicator of the direct-deflection type. Power is supplied to the receiver from the MCT secondaries connected series opposing. Transmission may be directly by alternating current or can be direct current if a suitable rectifier and filter is provided on the MCT output. A typical receiver indicator is a 0-to-10 volt, 1,000-ohm-per-volt voltmeter calibrated in terms of the measured variable.

A-C NULL-BALANCE SERVO SYSTEMS

A bridge-balance system using a resistance-ratio type of receiver is shown in Fig. 4. In this system the secondaries of

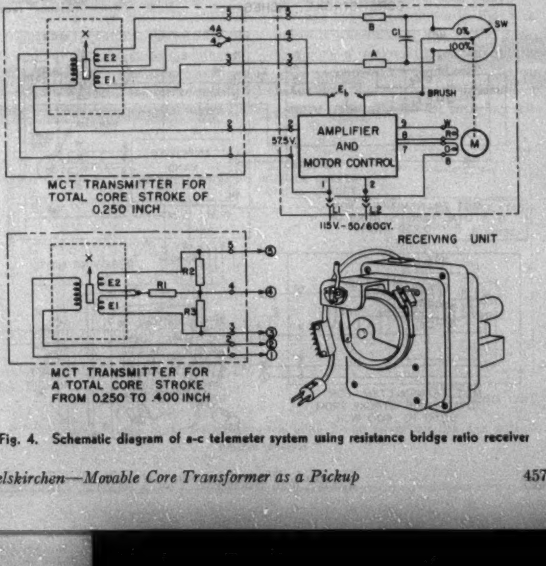

X
E2
E1
5
4A
4
3
2
1
B
C1
A
0%
100%
SW
BRUSH
E_b
57.5V.
AMPLIFIER AND MOTOR CONTROL
9
8
7
6
M
L1
L2
115V.-50/60CY.
MCT TRANSMITTER FOR TOTAL CORE STROKE OF 0.250 INCH
RECEIVING UNIT
R1
R2
R3
MCT TRANSMITTER FOR A TOTAL CORE STROKE FROM 0.250 TO .400 INCH

Fig. 4. Schematic diagram of a-c telemeter system using resistance bridge ratio receiver

the MCT are connected in series aiding and the ratio of induced secondary voltages is balanced by the ratio of the voltage drops across the resistance network in the receiver. When the telemeter bridge is balanced, the fundamental in-phase signal to the amplifier is zero and the following equation applies[7]

$$E_2/E_1=(B+S\theta)/[A+(1-\theta)S] \qquad (7)$$

where θ is the slide-wire motion expressed in per-unit terms as the ratio to total rotation.

Capacitor C is placed in the slide-wire circuit to compensate for the phase displacement that occurs between the two secondary voltages when the core is displaced. This manifests itself as a residual voltage at the amplifier input when the bridge is balanced. Since the in-phase component of the error signal is zero at balance, this residual voltage is quadrature and may saturate the amplifier, making the system appear to have less sensitivity. This is particularly evident at the slide-wire end points where this phase displacement between the MCT winding voltages is the greatest. Fig. 5 shows the residual voltage plotted as a function of core position over the entire range of core stroke with and without the capacitor C. The figure shows that the residual voltage is a minimum when the core is at the electrical center of the MCT where the winding phase displacement is the least.

On systems requiring a core stroke of equal displacement on either side of the MCT center, the compensating capacitor is connected across the entire slide wire. If the system calibration requires that the core start at the MCT center, the capacitor is connected from the contact brush to the zero end of the slide wire.

The standard ratio bridge receiver is calibrated for a 0.250-inch core stroke of the transmitter MCT core travel, which is adjusted by mechanical means for transmitters with linkage adjustments. In the integral type of transmitters for differential pressure and flow measurements, the MCT core is operated in a pressure-sealed tube and no linkage adjustments are available. For these applications, the resistance network shown in the lower portion of Fig. 4 converts the actual MCT core stroke to an equivalent 0.250-inch stroke for the standard receiver. The calibration is made by the adjustment of resistance R_1.

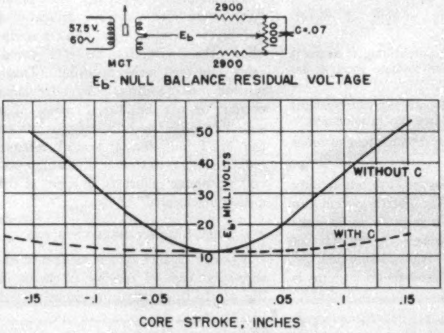
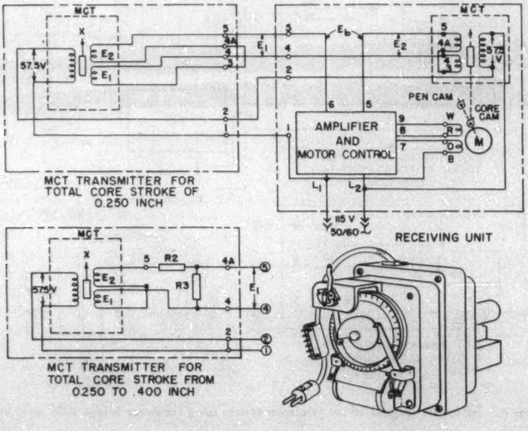


Fig. 5 (left). Graph of null-balance residual voltage for the a-c ratio bridge telemeter

Fig. 6 (below). Schematic diagram of a-c telemeter system using potential balance receiver


MCT
X
57.5V
E_2
E_1
5
4A
4
3
2
1
E_1
E_b
E_2
57.5 V
PEN CAM
CORE CAM
6
5
AMPLIFIER AND MOTOR CONTROL
9
8
7
W
R
O
B
M
L_1
L_2
115 V 50/60
MCT TRANSMITTER FOR TOTAL CORE STROKE OF 0.250 INCH
RECEIVING UNIT
R2
R3
⑤
④
②
①
MCT TRANSMITTER FOR TOTAL CORE STROKE FROM 0.250 TO .400 INCH

An a-c potential balance telemeter system using an MCT in the receiver is shown in Fig. 6. In this system the respective MCT secondaries are connected in series opposing so that the circuit is balanced when the MCT output voltages are equal and the fundamental in-phase signal to the amplifier is zero. The MCT primary windings are connected in series to equalize the primary magnetizing currents. The potential balance telemeter can be operated over a greater distance than the ratio bridge telemeter since there is no current flow in the secondaries when the circuit is in balance. The potential balance receiver is also calibrated for a standard 0.250-inch core stroke. For applications requiring a greater MCT core stroke, the voltage divider network shown in the lower portion of Fig. 6 is used to obtain the equivalent output of the standard stroke unit.

Both types of a-c null-balance systems perform satisfactorily with line voltage variations of 105 to 125 volts a-c and over a frequency range of 48 to 62 cycles. The ranges of operating voltage and frequency are established by the limitations of the servo amplifier and motor, since the MCT described in this paper can be excited by supply frequencies as high as 400 and 1,000 cycles and over a wide range of voltages. Other designs using powdered iron or magnetic ceramic cores and ceramic bobbins have been operated at frequencies as high as 100 kc.[5]

The standard receiver (for a 0.250-inch MCT core stroke) will respond to a 0.0001 inch change in core position and is accurate to 0.001 inch. However, motions in the order of 0.000010 inch can be detected with specially calibrated systems.

The a-c operated telemeter circuits are used for short-range applications and require special wiring precautions to minimize interference from adjacent power circuits. The ultimate distance limitation is determined by attenuation and phase shift in the transmitted signal produced by line resistance and interline capacity. However, the practical limit

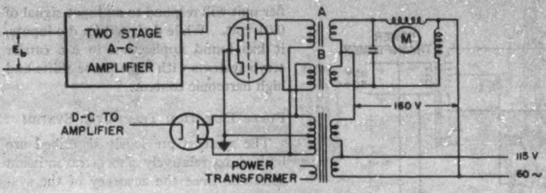


FIG 7(A) SATURABLE REACTOR MOTOR CONTROL

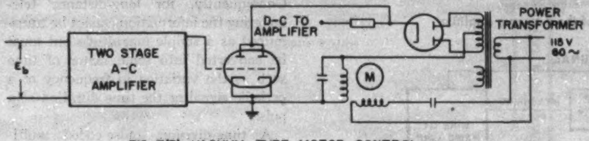


FIG 7(B) VACUUM TUBE MOTOR CONTROL

Fig. 7. Schematic diagram of a-c servo amplifiers for telemeter receivers

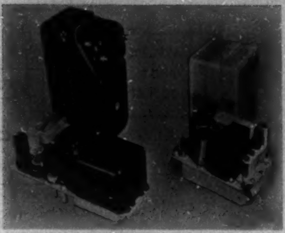


Fig. 8. A-c servo amplifiers

is established by the cost of shielded multiconductor cable as compared with the added cost of more complicated transmitter and receiver elements required for telephone, telegraph, or radio transmission.

A-C Servo Amplifiers

Fig. 7 illustrates two types of servo amplifier units which are being used in these a-c telemetering systems. Each of the units utilizes a high-gain resistance-capacitance coupled-voltage amplifier in addition to the motor control circuit for controlling the 2-phase balancing motor. Both circuits have about the same sensitivity and will respond to an error signal of 0.5 millivolt or less when the residual noise voltage of the system is less than 5 millivolts.

The circuit shown in Fig. 7(A) uses a pair of saturable reactors as the motor control elements. Differential d-c excitation of the reactor control windings controls the saturation to modulate the phase relation between the two motor winding currents. This system is relatively immune to harmonics and quadrature components in the amplifier error signal.

The circuit shown in Fig. 7(B) is designed for controlling a high-impedance servomotor with the control winding connected directly in the cathode circuit of the output tube. Damping of the motor is achieved by the d c component of the control winding current which is a maximum at balance. Fig. 8 shows the two types of a-c amplifiers and the considerable reduction in size achieved by the all-electronic unit using miniature tubes.

A magnetic amplifier has also been developed for these a-c null-balance systems where extreme reliability is required. Its size and performance are comparable to the electronic types described, but a detailed description is beyond the scope of this paper.

D-C Null-Balance Servo System

A d-c potentiometer telemeter for use with the MCT pickup is shown in Fig. 9. The advantage of zero-current flow at balance and the use of d-c signal for transmission eliminates effects of line-resistance phase shift and a-c pickup, and makes this system suitable for transmission distances up to 50 miles. The nominal signal used is 20 volts but the system will operate satisfactorily over a range as low as 5 volts.

A signal-converter unit, installed at the transmitter, contains the rectifier circuits and supplies a regulated voltage to the MCT primary. Secondary voltages from the MCT pickup are bridge-rectified and the d-c components appear across the individual load resistors *R*. The d-c output voltage of the signal converter is the algebraic sum of these components. Since the a-c components in each load resistance tend to cancel, the output ripple is kept at less than 3 per cent with single section filters. The accuracy and stability of the signal converter output depends upon the regulation of the MCT primary voltage and the stability of the rectifiers. The use of rectifiers having a very high inverse to forward resistance ratio minimizes the dependence on rectifier characteristics.

A conventional type of null-balance d-c potentiometer is used as the receiver, although any high-impedance d-c measuring device may be used. This receiver has a d-c servo amplifier to be described later and is useful for measurements of ranges of 5 d-c volts or greater without the use of choppers or other types of d-c to a-c converters. An electronic-regulated d-c supply furnishes the reference potential for the receiver slide wire.

While the system shown in Fig. 9 requires a transmission link of the 2-wire telephone or telegraph type of circuit, the MCT pickup with a signal converter can be used as a displacement to a direct-voltage-transducer for use with any tele-

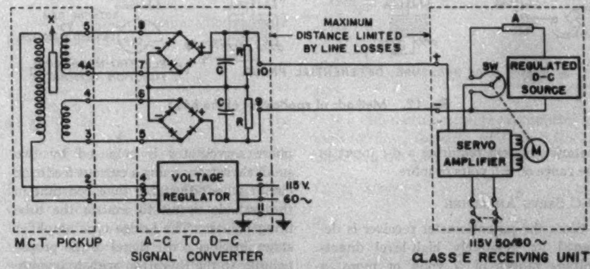


Fig. 9. Schematic diagram of d-c telemeter system using 2-wire telephone link and potentiometer receiver

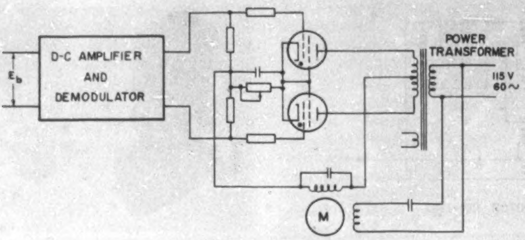

D-C AMPLIFIER AND DEMODULATOR
E_b
POWER TRANSFORMER
115 V 60~
M

Fig. 10. Schematic diagram of the d-c servo amplifier

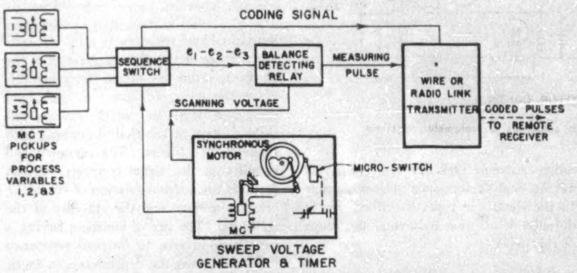

CODING SIGNAL
1
2
3
SEQUENCE SWITCH
$e_1 - e_2 - e_3$
BALANCE DETECTING RELAY
MEASURING PULSE
WIRE OR RADIO LINK TRANSMITTER
CODED PULSES TO REMOTE RECEIVER
SCANNING VOLTAGE
MCT PICKUPS FOR PROCESS VARIABLES 1, 2, & 3
SYNCHRONOUS MOTOR
MICRO-SWITCH
MCT
SWEEP VOLTAGE GENERATOR & TIMER

Fig. 11. Pulse modulation telemeter using scanning-type analogue to pulse duration converter

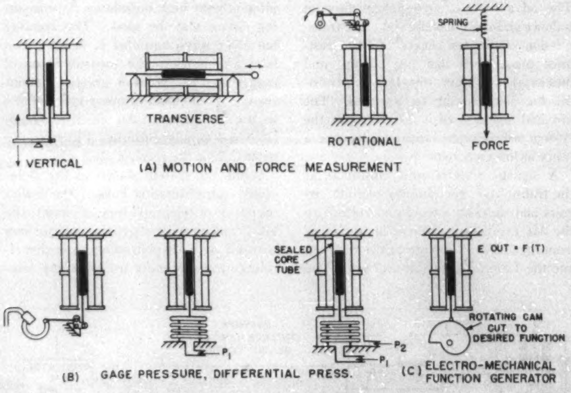

TRANSVERSE
VERTICAL
ROTATIONAL
SPRING
FORCE
(A) MOTION AND FORCE MEASUREMENT
SEALED CORE TUBE
E OUT = F (T)
P_1
P_2
P_1
ROTATING CAM CUT TO DESIRED FUNCTION
(B) GAGE PRESSURE, DIFFERENTIAL PRESS.
(C) ELECTRO-MECHANICAL FUNCTION GENERATOR

Fig. 12. Methods of application of the MCT

metering system requiring a d-c input in the range of ±5 volts or more.

D-C Servo Amplifier

Since the potentiometer receiver is designed for relatively high-level direct-voltage ranges of 5 volts or more, a simple electronic servo amplifier Fig. 10, is used. Proportional control of the 2-phase servomotor is achieved by two small thyratrons using a current feedback circuit with adjustable time constant in the cathode circuit to control the tube firing points. The bridge-type amplifier stage provides a d-c signal of the proper polarity to the thyratron grids to operate the servomotor in the direction to balance the potentiometer circuit. This amplifier unit will respond to an input signal of 0.01 volt. While designed for d-c inputs, it has found application in a-c carrier servo systems with large phase shifts and high harmonic content.

Pulse Duration Telemeter System

The systems previously described are limited to relatively short transmission distances since the accuracy of the systems is directly affected by attenuation of the telemetered signal amplitude. Consequently, for long-distance telemetering the information cannot be transmitted as a simple magnitude but must be converted into some factor of time such as the variation in frequency of a periodic signal or the time duration of a pulse.

A time-division, pulse-coded multiplexing system[8] meeting the requirements of a long-distance telemetering system is shown in Fig. 11. In this system, a number of process variables are translated into electric quantities, compared one at a time in sequence against a common scanning quantity, thus converting the magnitude of the transducer output into a proportional time pulse. A code selector pulse provides identification of each process variable.

The MCT is the transducer used to convert the displacement variables into a voltage magnitude. The output from each MCT transducer is selected by the sequence switch at the beginning of a cycle and compared against the scanner output by means of the balance-detecting relay. This balance-detecting relay allows a coded pulse to pass from the beginning of the cycle until the scanner voltage becomes equal to the transducer voltage, at which instant the coded pulse is stopped. Hence, the time duration of the coded pulse is made proportional to the voltage magnitude from the transducer. The sequence switch is then stepped to the next process variable and the cycle is repeated with a new coded pulse to identify the second measurand. The transmission link can be any medium, telephone or telegraph wire, carrier current system, radio, or any combination of these, once the process information has been converted into signals of time interval. The electromechanical scanning system shown in Fig. 11 is used where the scanning period can be several seconds. Distance is limited only by the transmission medium.

Application

The range of application of the MCT as the transducer element in mechanical dis-

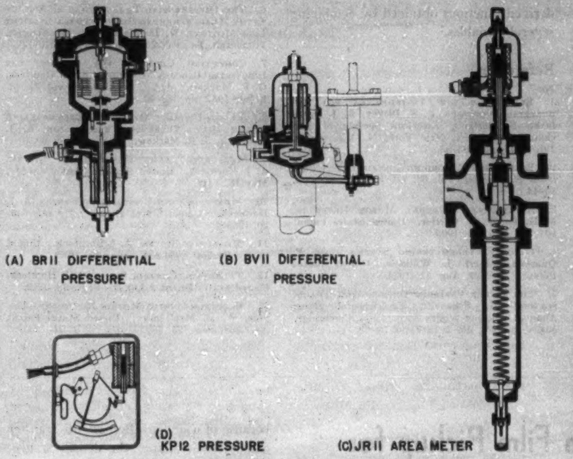
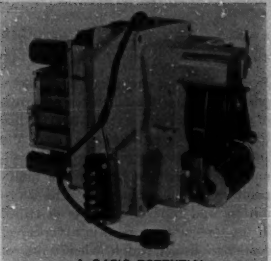
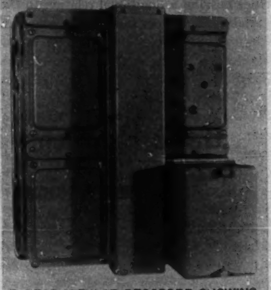

(A) BRII DIFFERENTIAL PRESSURE
(B) BVII DIFFERENTIAL PRESSURE
(D) KP12 PRESSURE
(C) JRII AREA METER

Fig. 13. Typical pickups for process quantities



A-BASIC POTENTIAL BALANCE RECEIVING UNIT



B-REAR VIEW OF RECORDER SHOWING RECEIVING UNIT MOUNTED IN PLACE

Fig. 14. Recording-type receiver

ol. 74
. 1
955

placement type of transmitters is extremely wide. While direct electric sensing elements are commonly used for certain process measurands such as temperature, gas analysis, density, and viscosity, many of the most important variables such as pressure, level, and flow are preferably measured by mechanical devices such as bellows, bourdon tubes, and the like. Fig. 12 shows in simple schematic form a number of typical methods of application of the MCT as the motion-to-electric transducer for physical quantities convertible to displacement.

Fig. 13 shows in more detail the actual design of several types of MCT pickups for pressure, differential pressure, and flow. The differential pressure and flow pickups shown in Figs. 13(A), (B), and (C) use a pressure-sealed, integral-core drive mechanism which results in a very simple, rugged transmitter requiring no pressure-sealed bearings. The output of these units is adjusted to the equivalent of 0.250-inch motion by means of the resistor network described previously. The pressure transmitter shown in Fig. 13(D) employs an MCT actuated by a bourdon tube through mechanical linkage. The stroke of the core is adjusted to precisely 0.250 inch so that the electric output is the same for all pressure ranges from 20 to 2,000 pounds per square inch full scale. The device is designed with an explosion-proof coil housing and is ambient-temperature-compensated to ±1/4 per cent from 40 to 140 degrees Fahrenheit.

A receiver element of the a-c null-balance servo type is shown in Fig. 14. This is designed as a complete package assembly with the servo amplifier arranged as a plug-in unit for ease of replacement for servicing. The drive link is actuated from a functional cam driven by the pickoff shaft. The figure shows the voltage-balance type of receiver using the MCT as the pickoff element. For maximum flexibility, to permit recording any function of the measurand, two cams are used. One positions the core of the MCT and the other the recorder drive arm as shown. The per-unit core motion x can be expressed as a function of M, the per-unit measurand

$$x=f(M) \tag{8}$$

At balance, the receiver core motion equals the transmitter core motion

$$y=x=f(M) \tag{9}$$

The cam which positions the receiver core is shaped so that

$$y=f(\theta) \tag{10}$$

hence

$$\theta=M \tag{11}$$

The slide-wire type of receiver used in the a-c resistance-ratio-balanced system is similar in construction to the unit just described except a slide wire replaces the MCT as the follow-up element. A functionally tapered slide wire is used with nonlinear measuring element or where a function of the measured quantity is to be recorded. The general balance equation between the per-unit slide-wire rotation θ and the per-unit core motion x is

$$1/S\int_0^\theta f(\theta)d\theta=x \tag{12}$$

where

$f(\theta)$ = the resistance change per-unit rotation
x = per-unit core motion
S = total slide-wire resistance

For linear recording using a uniform slide wire, $f(\theta)=S$ in the per-unit system.

Hence

$$1/S\int_0^\theta S d\theta=x \tag{13}$$

and

$$\theta=x \tag{14}$$

If a differential pressure transmitter is used for flow, the slide wire has a linear taper and the balance equation becomes

$$1/S\int_0^\theta 2S\theta d\theta=x \tag{15}$$

or

$$\theta^2=x \tag{16}$$

and

$$\theta=\sqrt{x} \tag{17}$$

Since

$$x = \text{flow}^2 \qquad (18)$$

$$\theta = \text{flow} \qquad (19)$$

In practice, it is customary to extract a part of the function on the cam because of the practical impossibility of extending the slide wire to zero width. Fig. 14 also shows the method of mounting of individual receiver units on a complete multipoint recording instrument. The transmitters and receiver elements described in the foregoing are used in a variety of combinations for simple telemetering of measured quantities, for the conversion of these quantities into secondary quantities bearing a functional relation to the variable, and for the computation of derived functions obtained by combining several variables.[7]

References

1. Electronic-Type Instruments for Industrial Processes, **P. S. Dickey, A. J. Hornfeck.** *Transactions*, American Society of Mechanical Engineers, New York, N. Y., vol. 67, July 1945, pp. 393–98.

2. Electronic Recording Instruments. *Electrical Engineering*, vol. 66, no. 1, Jan. 1947, pp. 36–44.

3. Alternating Current Motor Reverser, **E. H. Porter, B. Currier.** United States Patent *808,944*, Mar. 29, 1905.

4. Electric Telemetering System, **John F. Quereau, Albert J. Williams.** United States Patent *2,050,629*, Aug. 11, 1936.

5. The Linear Variable Differential Transformer, **H. Schaevitz.** Experimental Stress Analysis, Addison-Wesley Corporation, Cambridge, Mass., vol. IV, no. 2, 1947, pp. 79–88.

6. The Differential Transformer as Applied to the Measurement of Substantially Straight Line Motions, **W. D. MacGeorge.** *Instruments*, Pittsburgh, Pa., vol. 23, June 1950, pp. 610–14.

7. Computing Circuits and Devices for Industrial Process Functions, **A. J. Hornfeck, G. R. Markow.** *AIEE Transactions*, vol. 71, pt. I, July 1952, pp. 183–93.

8. A Long-Distance Multipoint Telemetering System Using Teletype Transmission, **A. J. Hornfeck, G. R. Markow.** *Ibid.*, pp. 193–99.

9. Measuring and Controlling System, **A. J. Hornfeck.** United States Patent *2,420,539*, May 13, 1947.

10. Measuring and Controlling System, **A. J. Hornfeck.** United States Patent *2,439,891*, Apr. 20, 1948.

11. Measuring System, **A. J. Hornfeck.** United States Patent *2,406,221*, Aug. 20, 1946.

12. Telemetric Control System, **A. J. Hornfeck.** United States Patent *2,330,427*, Sept. 28, 1943.

13. Electromagnetic Motion Responsive Device, **W. D. MacGeorge.** United States Patent *2,427,866*, Sept. 23, 1947.

Basic Problems in Film Pickup for Television Broadcasting

E. M. GORE
ASSOCIATE MEMBER AIEE

FILM programs now constitute a very large percentage of television (TV) air time. One network alone uses film at the rate of about 9,000,000 feet per month. Many of the major programs are recorded, for reasons of convenience, timing, or historical record. Some of the smaller stations have begun their operations using nothing but recorded material for all programs.

A minute portion of this enormous quantity of film has been used for the generation of color TV programs. Color programming is now steadily increasing and the volume of film used for color programs may in a few years equal the amount used for monochrome.

The use of film for TV involves a long chain of processes between the original scene and the reproduced image. Fig. 1 is a block diagram indicating the chain of processes involved. The shortcomings of all of these processes can accumulate, and satisfactory reproduction requires the solution of many technical problems. Evidence that the solutions are not unique is apparent from the different film camera equipments now being marketed. It is the aim of this paper to discuss some of the most important problems of TV film reproduction, especially from the standpoint of the TV system. A commercial color film camera system using three vidicon pickup tubes is described.

Film Characteristics

Film used for TV is produced by either photographing the original scene or by recording TV pictures. Positive prints are generally employed for monochrome use except that the original negative Kine-recording is frequently used for delayed broadcasts.

Color film requires much more care in photography and processing because of the interaction between the three dye images on the same film. Technical developments indicate that the need for color TV recordings can be met in the near future.[1]

The characteristics of film which are the most troublesome in TV reproduction because of wide variation are the transfer characteristic shape, the density range, and the color quality. Noise (graininess) and lack of sharpness are frequently very objectionable in 16-millimeter (mm) prints.

Improvement of TV film programs thus requires the co-operation of film manufacturers, photographers, and processors, in addition to the effort expended on the pickup and reproducing devices.

Film Camera Types and Television Projectors

Film cameras may be classified into two major types:

1. Flying-light-spot cameras focus a moving scanning spot, generated by a kinescope, onto the film frame. The transmitted light is proportional to the transmission of the film frame at that point and it is collected by a multiplier photocell to produce an electrical video signal.

2. Storage-type camera tubes accumulate a charge image of the focused optical image during the interval of light exposure, between scans by an electron beam. Storage-type tubes include the iconoscope, the image orthicon, and the vidicon.

Systems which are now being used in the United States for film pickup are the flying-spot scanner, the vidicon, and, of diminishing importance, the iconoscope and the image orthicon.

Both 16mm and 35mm films are shown at an average rate of 24 frames per second. United States TV operates at a rate of 30 frames or 60 fields per second. Thus a one-to-one correspondence between film frames and TV frames does not exist. The methods for conversion fall into two categories, intermittent projectors and continuous motion projectors.[2]

Paper 55-457, recommended by the AIEE TV and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 31, 1955; made available for printing April 19, 1955.

E. M. Gore is with the Radio Corporation of America, Camden, N. J.

The author expresses his appreciation for the information supplied by his associates in the Broadcast Television Equipment Group of the RCA Victor Division, and for their constructive criticism. Those most closely concerned were H. N. Kozanowski, S. L. Bendell, T. J. Shipferling, and N. L. Hobson.

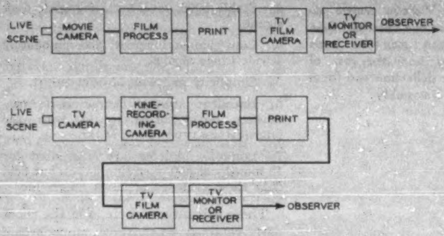


Fig. 1. Processes involved in TV reproduction of film

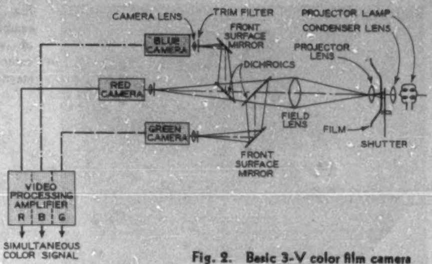


Fig. 2. Basic 3-V color film camera

Intermittent projectors utilize a shutter (or pulsed-arc lamp) to interrupt the light during pulldown or change of frames. These projectors may be subdivided into three types:

1. Projectors having uniform intervals between pulldowns allow 60 uniformly spaced exposures per second with short exposure time and moderately short pulldown time.

2. Projectors having nonuniform intervals between pulldowns, commonly called "3-2" mechanisms, expose alternate film frames three times and expose the remaining frames twice. Pulldown times are of the same order as those for theatre projectors.

3. Fast-pulldown projectors utilize the 3-2 mechanism and pull down the film within the interval of TV vertical blanking.[3]

Continuous motion projectors transport the film at uniform velocity. These projectors may be also classified into three types:

1. Stationary image projectors utilize moving lenses, prisms, or mirrors to maintain the projected image in a fixed position. The images of succeeding frames lap dissolve into each other.[1,4,5]

2. Single-line-scanning projectors use flying-spot scanning and depend on the film motion to give the vertical scan. These require special film operating at 60 frames per second.

3. Jump raster projectors utilize flying-spot rasters in two or more different positions for successive TV fields. This is accomplished by a step-centering wave form applied to the deflection coils of the flying-spot tube.[6]

With some combinations of projectors and cameras precise synchronism between film transport and TV scan rates is not required. This results in simplification and flexibility since the generation of motor drive power at the TV field frequency requires additional costly equipment.

For use with storage-type tubes, the 3-2 type of projector with moderate exposure is considered more reliable than fast pull-down projectors or continuous motion projectors. This type of projector is now being marketed in several commercial versions. Flying-spot systems using stationary image projectors are now being marketed.

Quality Characteristics of Film Cameras

Resolution may be expressed as equivalent line number or video bandwidth.[7] Since the receiver video passband is limited to about 4 megacycles per second, any of the cameras mentioned can provide adequate response.

Camera light requirements are determined by the minimum high-light brightness which will yield an acceptable signal-to-noise ratio. For the vidicon, however, the minimum high-light brightness is determined by the "lag" or long-time memory of the tube, which tendency decreases as the illumination is increased. In cameras where equalizers are employed to compensate for aperture responses the noise level is correspondingly increased.[8]

Transfer characteristic shape is linear for image orthicon and flying-spot pickup devices so that some form of electrical "gamma" correction is required to compensate for the nonlinear characteristic of the receiver kinescope. The iconoscope has variable gamma of less than unity while the vidicon has an essentially constant value of 0.65 for gamma. The useful brightness range of a camera is determined by the number of brightness steps of fixed ratio which can be identified in the electrical output signal. The maximum brightness input is usually limited by a saturation effect (as in the knee of the image orthicon) or by poor resolution due to the high beam current, or by the light source and optical system. The lowest brightness step discernible is usually determined by the electrical noise level. An undesirable behavior of some storage-type tubes, notably the iconoscope, is that the electrical signal level which represents black in the scene is a function of the average scene brightness.

Storage is the ability of a camera tube to retain a charge image after exposure to light and before scanning. With the intermittent type of projector an interval on the order of one half of a field may elapse between exposure and scan. In the order of increasing storage ability are the iconoscope, the image orthicon, and the vidicon.

Color response of the three storage-type camera tubes is panchromatic, although it is not identical to the standard visual response. By the use of suitable phosphors, photocells, and filters the flying-spot camera can be made panchromatic, although the red-channel sensitivity is usually much lower than the sensitivities of the blue and green channels.

Other Elements of Film Reproducing System

The observer's impression is the ultimate test of the job done by a TV system. The requirements of a TV reproducer are based on the "normal" response of observers with regard to luminance and color response.[9]

Luminance range and color sensation produced by the receiver kinescope may be affected by the amount and the color quality of the ambient light.[7,10,11] Both the ambient lighting conditions and the observer preference may vary widely. In the TV system it is therefore only the signal to the receiver which is standardized. These signal standards are based on a color kinescope transfer characteristic gradient (gamma) of about 2.2 and on a receiver color temperature of 6,500 degrees Kelvin when reproducing white.[11]

The signal transmission links between the camera chain and the receiver may reasonably be assumed to have a linear-transfer characteristic. In the color signal the chrominance information is

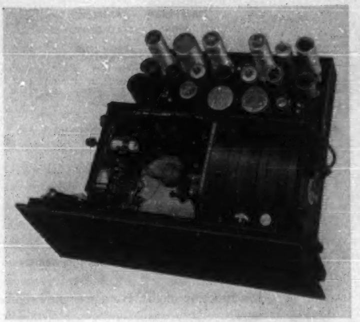


Fig. 3. Vidicon camera assembly for one channel. Base of vidicon may be seen extending from the rear of magnetic deflection and focus assembly

frequency multiplexed into the high-frequency portion of the video band occupied by the luminance information.

Corrective Devices and Methods

An obvious method for improving the TV reproduction of average color film is by improving the film itself. There are good color films available but many commercial color prints have very poor color fidelity and are quite grainy. Where films are produced exclusively for TV use the range should be limited to about 50 to one. The transfer characteristic should be of essentially constant gamma over this range. Tolerance limits on the color co-ordinates of certain commonplace colors might help avoid the often disagreeable appearance of flesh tones and of other familiar objects, even on direct projection.

Transfer characteristic gamma reduction by electrical methods increases the useful brightness input range for the camera, provided the camera noise level is sufficiently low. This method is particularly adapted to vidicon cameras because of the excellent signal-to-noise ratio which vidicon pickup tubes can provide. Electrical operations analogous to photographic masking techniques can be performed.[7,12,13]

The theoretical sensitivity curves for cameras contain negative lobes which cannot be realized optically. Further errors are introduced by the absence of negative lobes in the film taking sensitivities as well as by the dye crosstalk in the projected film. Partial correction can be achieved by the use of electrical masking or linear combinations of the three signals.[14] The optimum matrix is, unfortunately, not the same for all films.

An operation roughly equivalent to masking can be accomplished by modifying slightly with optical filters the shapes of the camera taking-sensitivity curves. Since the scene has already been analyzed into primary color components by the film taking sensitivities no great error should result from the analysis of the light transmitted by the film with narrow camera taking-sensitivity curves. The available signal with fixed illumination would be reduced and the signal-to-noise ratio would be decreased, but the degree of film dye crosstalk would be minimized.

Camera taking sensitivities can be so designed that one channel covers the entire visible spectrum and contains all of the luminance information while the remaining channels contain only chrominance information. These chrominance channels would operate with restricted bandwidth as required for the encoded signal. Monochrome receivers would utilize the information of only the luminance channel.

Aperture equalization permits compensation for some of the degradation of sharpness in the optical processes. The permissible degree of equalization is usually limited by the acceptable noise level so that it is important to minimize the degradations in each process.

Vidicon Color Film Camera System

The nonochrome vidicon film camera has replaced the iconoscope film camera in may installations during the past 3 years. Field experience has proved its ability to produce excellent pictures from good film and it is outstanding from the standpoint of reliability and cost of operation. The factors contributing to the good results are:[8,16]

1. Excellent signal-to-noise ratio, permitting a large amount of aperture equalization.
2. 600-line-limiting resolution. Approximately 40 per cent response at 350 lines without equalization.
3. Stable black level with negligible dark signal.
4. Low gamma inherent, accommodating a wide range of film.
5. Simplicity and ease of adjustment.
6. Excellent storage characteristic permitting nonsynchronous use with the intermittent type of projector.
7. Long-life vidicons are warranted for 12 months or 500 operating hours, but field data show evidence of much longer life.

These advantages were, for the most part, retained when three monochrome vidicon cameras were combined to form the 3-vidicon (3-V) color film camera.[16] A block diagram of this film pickup system is shown in Fig. 2. The light from the projector passes through a relay lens system and color separation filters where it is divided into three primary-color components, each of which is focused on one of the vidicon camera tubes. The video channels are independent so that the outputs are three simultaneous signals.

Because of the color separation filters the illumination required for the color film camera is nearly 20 times that required for the monochrome camera. Intermittent projectors having the 3-2 mechanism are used to obtain exposure cycles of about 50 per cent. By using a fast lens system adequate illumination is obtained for even the most dense color film. An optical multiplexer was designed to permit the use of a slide projector or one of two film projectors. The film projectors use 1-kw incandescent projection lamps with infrared filters so that still frames can be projected without damage to the film.

The output voltage of each of the three

Fig. 4. Projection room equipment for 3-V camera. Two 16mm TV projectors evident with optical multiplexer which reflects light into camera (foreground). Camera base houses deflection circuits and local controls

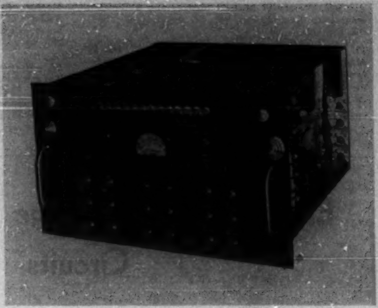


Fig. 6. Video processing amplifier. Four plug-in chassis used; mechanical design facilitates servicing

cameras, at any instant, must correspond to the transmission through a single picture element of the film frame. This registration requirement was fulfilled by the use of precision deflection and focus components, by sturdy mechanical construction, and by the use of stabilized deflection circuits. Fig. 3 shows the new camera assembly. Center resolution of 600 lines and corner resolution of 400 lines can be obtained on a superposition of the three signals.

Fig. 4 shows the equipment layout at the camera position. All operating controls are available at the remote-control desk. However, for convenience in setup these control functions are transferred to controls at the camera position by means of a switch. Vernier adjustments are provided on the registration controls.

Compensated video amplifiers on the top of each camera feed coaxial lines to the remote-control position. An additional output from each amplifier passes through a selector switch to a projection room monochrome monitor.

The control position is shown in Fig. 5. Camera-operating controls are located on the passive horizontal panel. The video signals from the camera enter the video processing amplifier shown in Fig. 6. This unit contains three parallel video channels which perform the operations of blanking modulation, gamma modification and shading (optional), and video power amplification. Functions performed in common for all channels are synchronizing pulse regeneration, monitor switching, and signal level calibration. This processing amplifier was developed as part of a project for color camera simplification. Considerable improvement in the state of the art is evident by the fact that the number of tubes in the image orthicon color camera chain was approximately halved as a result of this project.

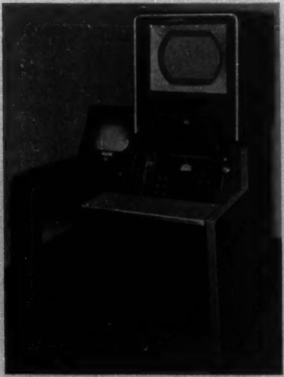


Fig. 5. Control position for 3-V color film camera. Front panel of video processing amplifier below color monitor. Remote-control panel for camera on desk (obscured) and monochrome master monitor with oscilloscope on left of processing amplifier

The remaining equipment, exclusive of power supplies, required to produce a National Television Systems Committee encoded signal consists of the colorplexer and its auxiliaries and a calibration pulse generator. Aperture equalization is provided in the luminance channel. The complete camera and encoding equipment uses five-power supplies providing approximately 2 amperes at 280 volt direct current regulated.

This 3-V color film camera has proved to yield very satisfactory film reproduction when used with good color film. Comparison of the resultant picture with a direct-projected picture shows remarkable fidelity. At present there are over 35 field installations in operation. The first installation, at The National Broadcasting Company in New York, has been in operation for more than a year.

Conclusions

A substantial improvement in the TV reproduction of color film can be achieved through better control of film, its exposure, and its processing, than is now evident in the average commercial color film. Receiver reproducing capabilities are limited as to contrast range, and consequently the film range should also be restricted.

Signal-to-noise ratio is a most important measurable quantity, applying both to the film and to the TV system. This ratio is basic to the determination of camera light requirements and contrast range, and may limit the corrections of gamma and detail rendition.

Certain electrical operations can be used to modify TV pictures in a manner similar to the photographic masking of films. Improvement of the reproduction of some films can be obtained by linear matrixing. Narrow TV camera taking-sensitivities can reduce dye cross-talk, but at the expense of signal-to-noise ratio.

The combination of the 3-V color film camera and intermittent projectors is capable of giving remarkable fidelity of color film reproduction. The low noise level of the vidicons permits operations such as aperture equalization and transfer characteristic modification while maintaining a good signal-to-noise ratio.

References

1. Color Transcription—Final Report of Panel 11A. National Television Systems Committee, Nov. 10, 1953.

2. Televised Film—A Review of Current and Proposed Methods, P. J. Herbst. *Broadcast News*, New York, N. Y., no. 69, May–June 1952, pp. 56–63.

3. Fast Cycling Intermittent for 16mm Film, Warren R. Isom. *Journal*, Society of Motion Picture and Television Engineers, Easton, Pa., vol. 62, no. 1, Jan. 1954, pp. 55–63.

4. New 35mm Television Film Scanner, E. H. Traub. *Ibid.*, pp. 45–54.

5. Continuous Motion Picture Projector for Use in Television Film Scanning, A. G. Jensen, R. E. Graham, C. F. Mattke. *Ibid.*, vol. 58, no. 1, Jan. 1952, pp. 1–21.

6. Continuous All-Electronic Scanner for 16mm Color Motion-Picture Film, Victor Graziano, Kurt Schlesinger. *Ibid.*, vol. 62, no. 4, Apr. 1954, pp. 294–305.

7. Image Gradation, Graininess, and Sharpness in Television and Motion Picture Images, O. H. Schade. *Ibid.*, Part I, vol. 56, no. 2, Feb. 1951, pp. 137–77; Part II, vol. 58, no. 3, Mar. 1952, pp. 181–222.

8. Vidicon for Film Pickup, R. G. Neuhauser. *Ibid.*, vol. 62, no. 2, Feb. 1954, pp. 142–52.

9. Color Television and Colorimetry, W. T. Wintringham. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 39, no. 10, Oct. 1951, pp. 1135–72.

10. Screen Viewing Factors Symposium.

Journal, Society of Motion Picture and Television Engineers, Easton, Pa., vol. 57, no. 3, Sept. 1951, pp. 185–224.

11. Color Balance for Television, **D. L. MacAdam.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 43, no. 1, Jan. 1955, pp. 11–14.

12. Electrical and Photographic Compensation in Television Film Reproduction, **P. J. Herbst, R. O. Drew, S. W. Johnson.** *Journal*, Society of Motion Picture and Television Engineers, Easton, Pa., vol. 57, no. 4, Oct. 1951, pp. 289–307.

13. Brightness Modification Proposals for Television Color Film, **W. L. Brewer, J. H. Ladd, J. E. Pinney.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 42, no. 1, Jan. 1954, pp. 174–91.

14. The Use of Electronic Masking in Color Television, **R. P. Burr.** *Ibid.*, pp. 192–200.

15. Vidicon Film Reproduction Cameras, **H. N. Kozanowski.** *Journal*, Society of Motion Picture and Television Engineers, Easton, Pa., vol. 62, no. 2, Feb. 1954, pp. 153–62.

16. 3-Vidicon Color Film Camera, **H. N. Kozanowski.** *Broadcast News*, New York, N. Y., no. 79, May-June 1954, pp. 8–15.

Engineering Multistage Diode Logic Circuits

B. J. YOKELSON
ASSOCIATE MEMBER AIEE

W. ULRICH
NONMEMBER AIEE

THIS article presents the method used for engineering the multistage diode logical function circuitry for a special-purpose computer employing exclusively semiconductor devices. Although the techniques used have been largely influenced by the devices available or being developed for the rest of the system, it is felt that these devices are typical of the general type of transistor device which may be expected in the near future. The methods to be described are partly graphical: no assumption of constant forward diode resistance is made. In this way, they differ from analysis methods based on Thévenin's theorem.

Reasons for Using Multistage Diode Logic

A logic circuit is defined as a collection of coincidence and mixing circuits, AND and OR gates, connecting a group of input devices with a group of output devices. A multistage diode logic circuit is one which uses only diodes as nonlinear elements, and one in which the output of one gate can serve as the input to a second gate, whose output in turn can serve as the input to a third gate, etc. In most computers, only two stages of logic are performed, and then a gain device of some sort is inserted. Since 2-stage logic is nonminimal, restriction to 2-stage logic involves a penalty: either gain is inserted to get a minimal diode network, or extra diodes are used and all logical operations operations are performed on a 2-stage basis.

The factors which normally restrict designers to 2-stage logic circuits are the following.

1. Time: multistage circuits introduce more delay than 2-stage circuits, if driven by similar input devices.
2. Nonuniformity of gates: multistage circuits are necessarily more nonuniform than 2-stage circuits.
3. Insufficient drive from pulsing devices: multistage circuits may require more pulse power than 2-stage circuits.

In the case of the computer for which these circuits were to be engineered, it was felt that the first two problems would not be critical, and since transistor switches were used as the pulsing devices, a large amount of pulse power was readily available. In general, it is felt that the advent of transistors will make multistage diode logic circuits practical.

Basic Diode Circuits for Performing Logic Functions

The two basic types of diode logic "building blocks" shown in Fig. 1 are well known. The circuits shown use negative pulses. In the AND circuit, the output is negative only if all inputs are negative, thus back-biasing all diodes. In the OR circuit, application of a negative potential at any of the inputs, biases that diode in the conducting direction and pulls the output negative.[1]

In this special-purpose computer, a separate signal was generated wherever the negative or prime of a primary signal was needed; this signal was the standard-pulse whenever its associated primary input was not pulsed, and was no-pulse whenever its associated primary input was pulsed. Thus, no inhibiting circuits were necessary except at the points where these prime signals were generated; with these exceptions, only logical AND and OR gates were used.

Design Techniques Used

General Remarks

In designing any circuit, the first consideration must be: What does the circuit do? In circuit language, this means: What are the output requirements? In logic circuits, the output requirements are determined by the sensitivity of the flip-flops, pulse amplifiers, and recording devices being fed by the logic. It is necessary to work from the output back to the input, and to work from the output requirements back through the logic and determine the input requirements.

Here, the first important difference between working with transistors and vacuum tubes becomes apparent: transistors are not infinite input impedance devices. The primary input considera-

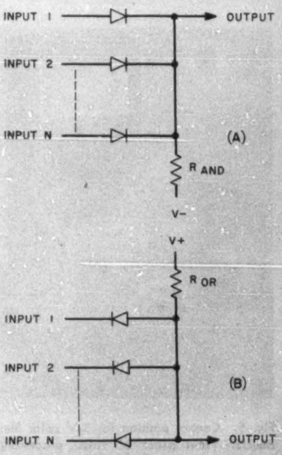


Fig. 1. A—Diode AND gate working with negative pulses. B—Diode OR gate working with negative pulses

---

Paper 55-431, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted February 8, 1955; made available for printing April 1, 1955.

B. J. Yokelson is with the Bell Telephone Laboratories, Inc., New York, N. Y., and W. Ulrich is with the Bell Telephone Laboratories, Inc., Whippany, N. J.

tion is current; even relatively sensitive devices require on the order of 1 milliampere (ma) to trip them. Thus, the impedance levels are lower than those encountered in vacuum-tube logic circuits. It is therefore expected that capacitance problems will be less serious than those for equivalent vacuum-tube circuits.

IDEAL COMPONENT THEORY

It is a common misconception that with ideal, continuously variable elements lossless diode AND and/or OR circuits can be obtained. This is not the case with the simple types of gates discussed in this article. To develop the design equations, ideal components will first be assumed and the corrections for imperfect diodes, resistor tolerances, and voltage regulation will be inserted in the equations as correction terms.

The following conventions on notation will be used; see Fig. 2. All pulses are negative, a pulse being the active condition, and no pulse being passive:

V_+ = positive power supply voltage
V_- = negative power supply voltage

Either $V+$ or $V-$ may be ground; intermediate voltages may be derived by means of a voltage divider network.

$V_{1\,\text{out}}$ = negative limit of the positive or passive truth voltage at the output of a gate

"Positive" and "negative" are used here as relative terms only. Thus, a positive truth voltage may be negative with respect to ground, but will certainly be more positive than the negative truth voltage. Similarly, positive limit means upper limit in the algebraic sense, negative limit means lower limit in the algebraic sense.

$V_{0\,\text{out}}$ = positive limit of the negative or active truth voltage at the output of a gate
$V_{1\,\text{in}}$ = negative limit of the positive or passive truth voltage at the input to a gate
$V_{0\,\text{in}}$ = positive limit of the negative or active truth voltage at the input to a gate

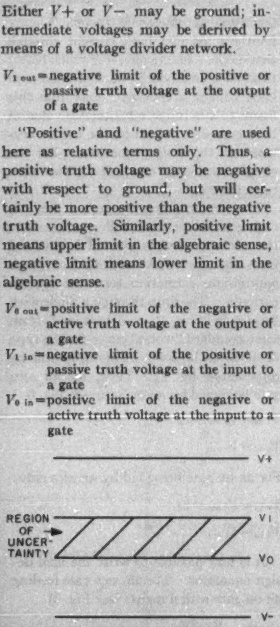


Fig. 2. Relative voltage levels of power supply voltages, active and passive voltage limits, and region of uncertainty

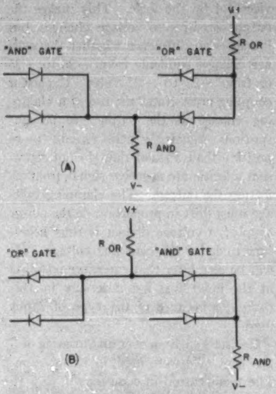


Fig. 3. A—Diode AND gate feeding diode OR gate. B—Diode OR gate feeding diode AND gate

With ideal diodes, the input and output voltages of a gate are the same; with actual diodes, the forward resistance of the diode must be taken into account.

The region of uncertainty is a band of voltages which should be avoided; if the output voltage of a gate is in this region, it cannot be guaranteed that the output device will not trip or that it will trip.

AND *Gate Feeding* OR *Gate*

Suppose there is a circuit where an AND gate is one of the inputs to an OR gate; see Fig. 3(A). The requirement is that when all the inputs to the AND gate are negative, the AND gate resistor pulls the output of the OR gate negative. Thus, the output current of the AND gate must be sufficiently high so that if all of it flows through the OR gate resistor (with ideal diodes, all the current from the AND gate will flow to the OR gate resistor) the output of the OR gate becomes negative by the required amount.

Thus, if the OR gate resistor is R_{OR}

$$I_{\text{out min}}\ \text{required} = \frac{V_+ - V_0}{R_{OR}} \qquad (1)$$

With ideal diodes, the inputs to the AND gate are all effectively shut off; therefore, the maximum possible output current with a given AND gate resistor R_{AND} is

$$I_{\text{out max}} = \frac{V_0 - V_-}{R_{AND}} \qquad (2)$$

An optimum circuit will have R_{AND} as large as possible so that the input pulse current is a minimum. The maximum R_{AND} must still deliver $I_{\text{out min}}$ required. Therefore, by equating equations 1 and 2, the maximum value of R_{AND} for a given R_{OR} can be obtained

$$R_{AND\,\text{max}} = R_{OR}\,\frac{(V_0 - V_-)}{(V_+ - V_0)} \qquad (3)$$

OR *Gate Feeding* AND *Gate*

Suppose an OR gate feeds an AND gate; see Fig. 3(B). The requirement is that if all other inputs to the AND gate are negative, and if all inputs to the OR gate are positive, the OR gate resistor supply enough current to the AND gate to keep the output of the AND gate positive. If the AND gate resistor is R_{AND}

$$I_{\text{out min}} = \frac{V_1 - V_-}{R_{AND}} \qquad (4)$$

The maximum output of the OR gate with resistor R_{OR} is

$$I_{\text{out max}} = \frac{V_+ - V_1}{R_{OR}} \qquad (5)$$

Using similar reasoning, $R_{OR\,\text{max}}$ can be obtained by equating equations 4 and 5; therefore

$$R_{OR\,\text{max}} = R_{AND}\,\frac{(V_+ - V_1)}{(V_1 - V_-)} \qquad (6)$$

It is now possible to solve for the ratio of an AND gate resistor which feeds an OR gate, to the AND gate resistor fed by the same OR gate; see Fig. 4. Let the AND gate resistor feeding the OR gate be R_1 and the AND gate resistor being fed by the OR gate resistor R_2.

With the use of equation 6

$$R_{OR} = R_2\,\frac{(V_+ - V_1)}{(V_1 - V_-)} \qquad (7)$$

With the use of equation 3

$$R_1 = R_{OR}\,\frac{(V_0 - V_-)}{(V_+ - V_0)} \qquad (8)$$

If equations 7 and 8 are substituted

$$R_1 = R_2\,\frac{(V_+ - V_1)}{(V_1 - V_-)}\,\frac{(V_0 - V_-)}{(V_+ - V_0)}$$

$$\text{or}\ \frac{R_1}{R_2} = \frac{(V_+ - V_1)}{(V_+ - V_0)}\,\frac{(V_0 - V_-)}{(V_1 - V_-)} \qquad (9)$$

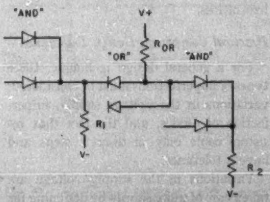


Fig. 4. Diode AND gate feeding OR gate feeding AND gate

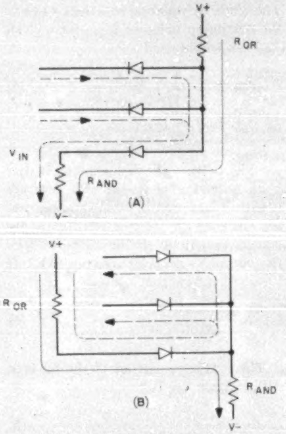


Fig. 5. A—Paths of current for an AND gate feeding an OR gate. B—Paths of current for an OR gate feeding an AND gate

Solid arrows indicate paths of main current
Dashed arrows indicate paths of reverse diode current

By inspection of the terms in equation 9 (see Fig. 2 for a visual description of the factors in the equation) it is obvious that R_1/R_2 is less than 1. Thus, input current at the first stage must be fed into a lower value resistor R_1 than input current at the third stage, even though in this ideal case, the input voltages are the same. This proves that with these types of diode gates, loss is not only a function of the imperfection of the components used, but that some loss is inherent. It can be seen that the percentage loss can be made as small as desired for a given output voltage swing by ever increasing the supply voltages and by placing V_1 and V_0 about halfway between V_+ and V_-; see Fig. 2. Practically, however, the problems introduced by increasing the supply voltages beyond a reasonable amount will be far greater than the problem of loss.

Practical Component Design Techniques

For a general design technique, three types of imperfections must be considered: variations in the voltage supply, imperfection of diodes, and the fact that resistors come only in discrete steps and have a tolerance.

Variations in the supply voltage are taken care of quite simply by designing for the minimum supply voltage condition, and using capacitor coupling to the devices fed by the logic. This makes the output sensitive to voltage changes, but it makes no difference whether the voltage changes suddenly from −20 to −25 or from −22 to −27 volts. Capacitor coupling necessitates the use of a clamping voltage at the output, to prevent spurious signals, which remain more positive than V_1 and thus do not represent a legitimate negative signal, from affecting the output. The clamping voltage must drift in proportion to the power supply; a voltage divider is thus necessary to derive the clamping voltage. (It will be seen later that capacitor coupling at the input was not necessary in this computer because of the type of input used.)

Resistors with an over-all variation of 5 per cent (%) were used, in this system. The modification in equations 1 to 6 to take care of the resistor variation, always assuming that the resistor variation is in a direction to decrease the output, results in equations 1(A) to 6(A). R_{AND}, R_{OR}, $R_{\text{AND max}}$, and $R_{\text{OR max}}$ all refer to nominal resistance values. Thus, $R_{\text{AND max}}$ is the maximum nominal value of the resistor while 1.05 $R_{\text{AND max}}$ is the maximum value.

$$I_{\text{out min}}=\frac{V_+-V_0}{(0.95)(R_{\text{OR}})} \tag{1A}$$

$$I_{\text{out max}}=\frac{V_0-V_-}{(1.05)(R_{\text{AND}})} \tag{2A}$$

Therefore

$$R_{\text{AND max}}=\frac{(0.95)(R_{\text{OR}})(V_0-V_-)}{(1.05)(V_+-V_0)} \tag{3A}$$

$$\text{and } I_{\text{out min}}=\frac{V_1-V_-}{(0.95)(R_{\text{AND}})} \tag{4A}$$

$$I_{\text{out max}}=\frac{V_+-V_1}{(1.05)(R_{\text{OR}})} \tag{5A}$$

Therefore

$$R_{\text{OR max}}=\frac{(0.95)(R_{\text{AND}})(V_+-V_1)}{(1.05)(V_1-V_-)} \tag{6A}$$

The maximum values of resistors with the use of these equations, having been found, the next lowest available standard values are used.

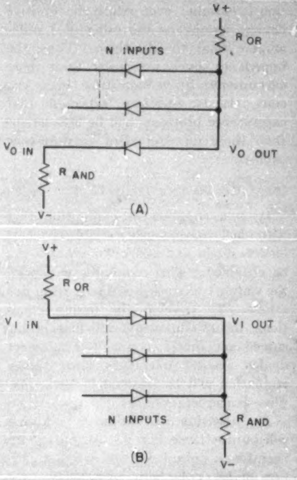


Fig. 6. A—Terms of equations 10 through 12. B—Terms of equations 13 through 15

These equations still do not take care of nonideal diodes. The AND gate resistor supplying an OR gate must supply any reverse current flowing to the other nonactivated inputs to the OR gate. Similarly, the OR gate resistor feeding an AND gate must supply all the back currents to the negative inputs to the AND gate; see Fig. 5. Theoretically, every diode position in the circuit should be examined individually to check on the maximum back voltage for this particular diode, to determine some reasonable figure for the maximum back current through that diode. Practically, it is much better to assume a reasonable value of maximum back voltage everywhere, and from this figure to determine a reasonable value of back current i_b, which can be used everywhere in the design equations. For the circuits used in the computer, i_b was taken to be 0.1 ma.

The forward drop through a diode is compensated for in the following manner: the minimum output current is obtained from equations 1(A) and 4(A), with an appropriate correction term for reverse currents. The drop through a diode with this current is then determined from some standard "worst" curve for the type of diode to be used. This drop will be represented by V_d. Then, for an AND gate being fed by an OR gate; see Fig. 6

$$V_{1\text{ in}}=V_{1\text{ out}}+V_d$$

$$V_{0\text{ in}}=V_{0\text{ out}}$$

For an OR gate being fed by an AND gate

$$V_{0\text{ in}}=V_{0\text{ out}}-V_d$$

$$V_{1\text{ in}}=V_{1\text{ out}}$$

It is now possible to write the final design equations. For an AND gate feeding an OR gate with n inputs (see Fig. 6)

$$I_{\text{out min}}=\frac{V_+-V_{0\text{ out}}}{(0.95)(R_{\text{OR}})}+(n-1)i_b \tag{10}$$

$$V_{0\text{ in}}=V_{0\text{ out}}-V_d \tag{11}$$

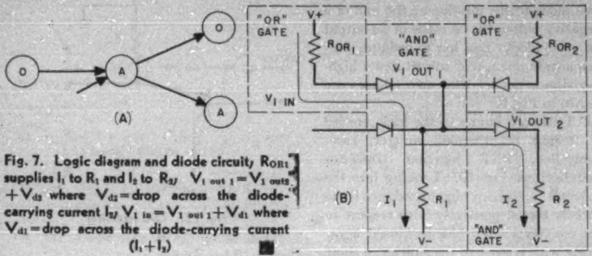


Fig. 7. Logic diagram and diode circuit; R_{OR1} supplies I_1 to R_1 and I_2 to R_2; $V_{1\,out\,1}=V_{1\,out\,2}+V_{d2}$ where V_{d2}=drop across the diode-carrying current I_2; $V_{1\,in}=V_{1\,out\,1}+V_{d1}$ where V_{d1}=drop across the diode-carrying current (I_1+I_2)

where V_d is the forward drop through a diode with $I_{out\,min}$ flowing

$$R_{AND\,max}=\frac{V_{0\,in}-V_-}{(1.05)(I_{out\,min})} \tag{12}$$

For an OR gate feeding an AND gate with n inputs

$$I_{out\,min}=\frac{V_{1\,out}-V_-}{(0.95)(R_{AND})}+(n-1)i_b \tag{13}$$

$$V_{1\,in}=V_{1\,out}+V_d \tag{14}$$

where V_d is the forward drop through a diode with $I_{out\,min}$ flowing

$$R_{OR\,max}=\frac{V_+-V_{1\,in}}{(1.05)(I_{out\,min})} \tag{15}$$

In the case where an AND gate feeds directly into an AND gate and an OR gate, the AND gate is designed as if it were feeding an OR gate only. The following precaution must be taken: since the AND gate has two output branches, the required V_0 at the output of the AND gate must be the more negative of the two input V_0's. If the input V_0 to the AND gate is more negative than the input V_0 to the OR gate, it is the more negative voltage which must be supplied to the OR gate. In this case the output current from the AND gate is calculated as if there were no drop across the diode to the OR gate since this gives a higher output current requirement. The inputs to this AND gate however, must be designed to feed both this AND gate and the AND gate following. Care must be taken to include two forward diode drops between the input and the second stage AND gate, the diode drops being different because different amounts of current flow through the two diodes; see Fig. 7. Similarly, where an OR gate feeds directly into an AND gate and another OR gate, the OR gate is designed to feed only the AND gate, and the inputs to the OR gate must feed both OR gates. (The same precaution as just cited is necessary. If V_1 to the OR gate is more positive than V_1 to the AND, the more positive V_1 must be supplied to the AND gate also. The output current from the OR gate is calculated as if there were no drop across the diode to the AND gate since this gives the maximum required output current.) In cases where an AND gate feeds only AND gates, or an OR gate feeds only OR gates, no resistor is needed.

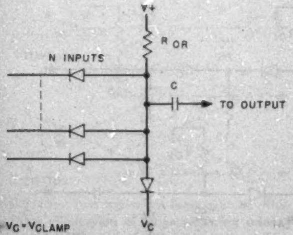


Fig. 8. Output OR gate

In the case where an OR gate feeds only an output and does not feed any other gates, the OR gate has only the function of restoring the d-c level on the logic side of the output condenser before the next signal occurs, Fig. 8, provided that only the leading edge of a pulse is used, and that a slowly decaying trailing edge can be tolerated. Since rather small coupling condensers are used, the value of the OR gate resistor is quite high, and relatively little extra loss is introduced by the output OR stage.

The maximum value of this resistor can be obtained easily if it is remembered that the net amount of charge flowing into the condenser from the logic must be zero over the shortest interval from the beginning of one pulse to the beginning of the next if small coupling condensers are used, and over a much longer interval if large condensers are used.

With a small coupling capacitor the problem can best be handled by solving separately for the charging and discharging condition. One simple approach is to assume linear charging, exponential discharging. Let t_1 be the pulsing time, t_2 the minimum time before the next pulse. (No provision is made here for condenser tolerance, which can be handled in the same way as resistor tolerance.) ΔV_{pulse} is the change of voltage across C while pulsing. Then

$$\Delta V_{pulse}=\frac{(t_1)(i_{pulse\,max})}{C}$$

provided this quantity is less than the magnitude of the pulse.

$$\frac{V_+-V_c}{V_+-V_c+\Delta V}=e\left(\frac{-t_2}{R_{OR\,max}C}\right)$$

Therefore, taking resistor tolerance into account

$$R_{OR\,max}=\frac{t_2}{(1.05C)\,1n\left(\frac{V_++\Delta V-V_c}{V_+-V_c}\right)}$$

$$R_{OR\,max}=\frac{t_2}{(1.05C)\,1n\left(1+\frac{t_1 i_{pulse\,max}}{C(V_+-V_c)}\right)} \tag{16}$$

If the magnitude of the pulse is less than $t_1(i_{pulse\,max})/C$, use V_{pulse} instead. In this case

$$R_{OR\,max}=\frac{t_2}{(1.05C)\,1n\left[1+\frac{V_{pulse}}{V+-V_c}\right]}$$

When an AND gate feeds an output OR

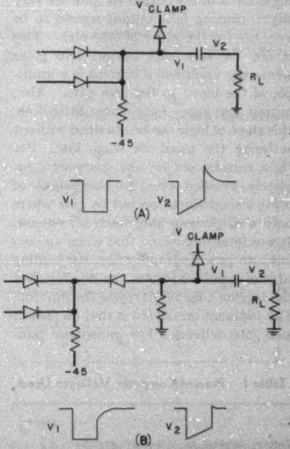


Fig. 9. A—Output AND gate with output voltage wave shape. B—Output OR gate with output voltage wave shape; illustrates elimination of overshoot

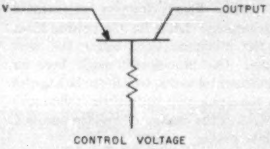


Fig. 10. PNP transistor switch

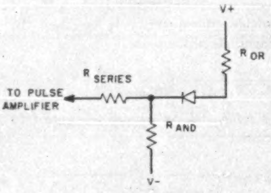


Fig. 11. Circuit for pulse amplifier feeding OR gate directly

gate only, the AND gate is designed in the usual manner except that the required current to the output is added to the term $I_{\mathrm{out\ min}}$. The output voltage requirements are applied to the output of the output OR gate, with the appropriate d-c shift so that the positive output is the same as the clamping voltage. Output AND gates are avoided because clamping poses a much more severe problem in the case of an output AND gate than with an output OR gate. Variations of clamping current in the output OR gate are very slight, causing false output signals to be very low (on the order of 0.2 volt). This is not the case with output AND gates where the variation of current is a multiple of the input to the AND gate. Also, output OR gates have very little loss; this stage of logic can be obtained without suffering the usual OR stage loss. For these reasons, output AND gates are completely avoided even to the extent of using a single-input output OR gate where only a single AND gate feeds an output. It was later discovered that when an output AND gate feeds a flip-flop, the leading edge of a pulse operates the flip-flop, but the lagging edge often resets the flip-flop. The explanation offered is that an output AND gate delivers a low impedance positive pulse to the flip-flop at the end of a negative pulse; the OR diode of an output OR gate blocks this low-impedance low-magnitude pulse, and substitutes a high-impedance pulse. The wave shapes are shown in Fig. 9.

Table I. Present Computer Voltages Used

V_+	ground
Voltage applied to transistor switch emitters	−12 volts
Positive clamping voltage	−20 volts
Positive limit of negative truth voltage	−25 volts
Negative voltage clamp of output of transistor switch	−30 volts
V_-	−45 volts

All primary inputs to the logic are fed by means of grounded emitter, base-controlled, PNP junction transistor switches; see Fig. 10. Looking into the collector, one sees a high impedance if the base is biased positively with respect to the emitter; if the base is biased negatively with respect to the emitter, a short circuit may be seen from emitter to collector. An AND gate requires current input when inactive, an OR gate requires current when active. Since more of the primary inputs to the logic circuits in this system are to AND gates (this is probably true in most logic circuits) a transistor switch which is pulsed open was used. This means that the collector of the switch can be connected directly, or through a dropping resistor, to the diode of an AND gate. When an OR gate is fed, a single input AND gate must be used to act as input to the OR gate, i.e., the AND gate will supply current to the OR gate; see Fig. 11. Initially, the emitters of the transistor switches were connected to V_+. It was recognized later that connection to some derived voltage between V_+ and V_1 would bring about the advantages of a lower reverse voltage across diodes and transistors, and would eliminate the necessity of supplying a voltage more positive than V_+ to the base of the PNP transistor switch when the transistor is pulsed off. The present voltages used in the computer are shown in Table I.

All these voltages are derived from voltage dividers or equivalent circuits; −45 volts is the positive limit of the battery (−48 volts nominal) excursions. The voltages are specified for the condition that the battery is at −45 volts; if the battery is at −48 volts, all voltages will increase proportionately, i.e., −12 goes to −12.8, −30 goes to −32, etc.

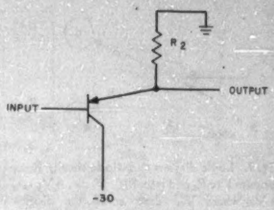


Fig. 12. D-c amplifier

Some Special Techniques

Some of the more specialized problems which may arise and some auxiliary devices will be discussed in the following. One such device is an emitter-follower d-c amplifier consisting of a single transistor; see Fig. 12. The negative input voltage must be about 1 volt more negative than the required output, the positive input being the same as the required positive output. The negative input current is $(V_{0\ \mathrm{out}}/R_2)\ (1-\alpha/\alpha)$. With α's of 0.95 or better, a large current gain may be obtained. These devices are useful when a very long chain of logic is encountered, or when many gates are fed from a single gate. Care must be taken if an OR gate feeds a d-c amplifier and the amplifier needs fairly rapid recovery, to provide a sufficiently low OR gate resistance. From a simple analysis no OR gate resistor is necessary but, actually, some means must be found to dispose of the minority carriers generated in the transistor. This matter must be explored experimentally.

A situation which is frequently encountered is that of one gate feeding a chain and a single gate; see Fig. 13. Although the voltage swing required by the single gate may be much less than that required by the chain, the feeding gate must supply the larger swing and, hence, the larger current to both gates since the input to both will be the same. (Strictly speaking, a resistance could be inserted

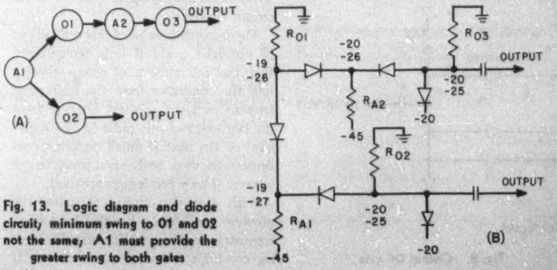


Fig. 13. Logic diagram and diode circuit; minimum swing to O1 and O2 not the same; A1 must provide the greater swing to both gates

Vol. 7
Pt. 1
1955

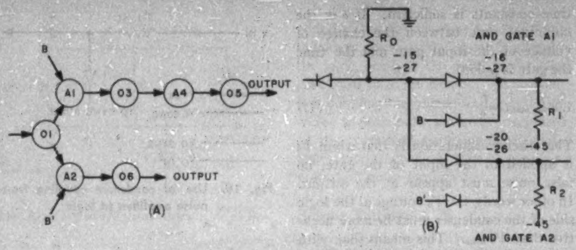


Fig. 14. Logic diagram and diode circuit; R_{OR} need not feed A1 and A2 simultaneously, since B and B′ cannot be pulsed simultaneously

between the feeding gate and the single gate, and so only the minimum current provided to the single gate; the advantage gained is much less than the cost of the resistor.) In Fig. 13, gate A_1 must supply a swing from −19 to −27 volts, even though gate o_2 needs a swing from only −20 to −26 volts.

Should a situation arise such as is shown in Fig. 14, where it is impossible that the other inputs to both AND gates be pulsed simultaneously, it is possible to assume that the drain from the OR gate to the AND gate of higher impedance is zero, provided that the input swing of the other input to the higher impedance gate, is larger than the input swing required from the OR gate. A moment's reflection will show that this is true. If the output swing from the OR gate is less than that from input B' to AND gate A_2, B' will supply all the power. If for some reason the output swing of the OR gate should increase beyond the swing of B', then the input to gate A_1 being the same as the input to gate A_2, will certainly be large enough to supply gate A_1. Thus, the OR resistor may be designed as if the OR gate fed only gate A_1, provided the foregoing precautions are taken. For example, in Fig. 14, by providing an input swing at B and B' of from −14 to −27, the heaviest drain on OR gate o_1 occurs when B is pulsed and A_1 is fed from o_1 alone. (R_1 is assumed to be less than R_2.)

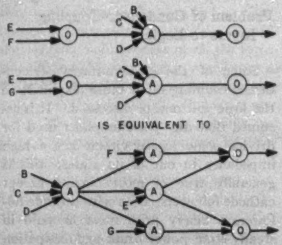


Fig. 15. Two equivalent logic circuits; illustrates elimination of OR gates having primary inputs

The transistor switches which feed the logic deliver more current when other transistor switches are pulsed off. Since the instantaneous maximum output current is limited by the transistor switch and its control, it is important to limit the peak current.

To get the most even distribution of current among all pulse amplifiers, it is desirable to increase their source impedance. A resistor is therefore usually placed in series with a pulser to prevent that pulser from delivering too much current, if it happens to have a very low internal impedance and is feeding many gates.

Input OR gates are undesirable, especially if a long chain of logic is involved. They introduce an extra stage of loss because of the dummy AND stage, Fig. 11, required when using the type of pulse amplifier described in the foregoing. If a pulsed-on type of transistor switch were used, dummy OR stages would have to precede AND gate inputs, while OR gates could be pulsed directly.

It is frequently possible to change the logic in such a way as to eliminate input OR gates (see Fig. 15 for an example). When this is not possible, an AND gate resistor must be calculated, capable of pulsing the OR gate; see Fig. 11. This resistor is connected either between the output of the pulse amplifier and the negative supply, or between a series resistor and the negative supply. This series resistor is calculated as indicated in the foregoing: if used, it will save input power and lower the back voltage on diodes. Theoretically, a diode may be necessary between the series resistor and the AND gate resistor, but practically it is not necessary since for reasons involving the circuitry of the pulse amplifier, it is desirable to have some resistance from the output of a pulse amplifier to $V-$. (To obtain a good rise time from this type of pulse amplifier, it is desirable to insure that at least 10 to 20% of the design peak load always be attached to the pulse amplifier.)

Frequently, relay contacts will act as inputs to logic circuits. Since a pulse amplifier acts as a switch, the obvious idea is to use the relay as an ideal switch in the place of a pulse amplifier. The equivalent of a pulse amplifier is therefore a break-contact with −12 volts connected on one side; see Fig. 16. If an OR gate is fed by an AND gate, one of whose inputs is a relay contact, this input can be replaced by a make-contact between the AND gate and $V-$; see Fig. 17. If, in addition, all inputs to this OR gate are AND gates, one of whose inputs is a relay contact, only one of the relays being operated at a time, the make-contact can replace the OR gate diode; see Fig. 18(A). In this case, the AND resistor is calculated in the usual manner, except that there is no voltage loss going through the relay contact. The circuit in Fig. 18(B) may be used if each of the AND gates in Fig. 18(A), has only one pulse input and one relay input, only one relay being operated at one time.

Transmission-Type Gates

A useful special-purpose gate is the transmission-type gate, in which the inputs are used to charge a condenser;

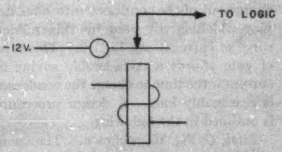


Fig. 16. Relay contact-feeding logic circuit

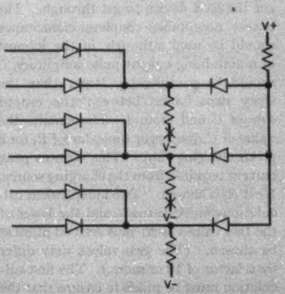


Fig. 17. OR gate fed by composite AND gate (diodes and relay contact) and other composite AND gates and/or diode AND gates; relays may be operated in any combination

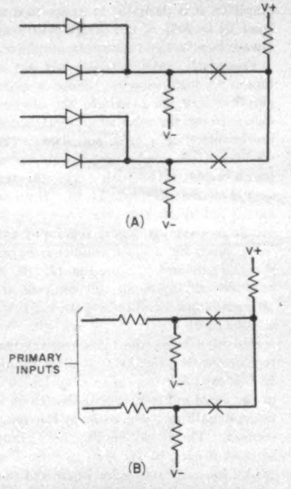


Fig. 18. A—OR gate fed by composite AND gates (diodes and relay contacts); relays are operated one or none at a time only. B—OR gate fed by composite AND gates (one primary input and relay contact) relays operated one or none at a time only

when the output is required, a pulse is applied in such a way as to spill the stored charge out of the condenser into some device which is triggered by this sudden surge of current; see Fig. 19. This type of gate affords a considerable saving in current if the time to charge the condenser is reasonably long. A design procedure is outlined in the following.

First, *C*, Fig. 19, is chosen. The value of the coupling capacitor is one which will allow a pulse of length necessary to trigger the load device to get through. The lowest acceptable coupling capacitance should be used although, in the interest of maintaining output pulse amplitude *C* should be considerably larger than the stray capacitance between the output side of *C* and ground. The smaller the value of *C*, the larger the value of R_1 for a given charging time. This reduces peak current required from the charging source. Next R_1 is chosen. Two independent calculations must be made and the lower of the two values of R_1 thus arrived at must be chosen. (The two values may differ by a factor of 10 or more.) The first calculation must be made to ensure that the condenser has time to charge. In transmission-type gates delivering negative pulses $V_{1\,\text{in}}$ and $V_{0\,\text{sync}}$ are active, $V_{0\,\text{in}}$ and $V_{1\,\text{sync}}$ are passive. Allowing three time constants is sufficient. If t is the minimum time between the change of voltage of the input gate and the time the gate is pulsed

$$1.05R_{1\,\max}C=\frac{t}{3} \tag{17}$$

The other consideration is that when V_0 is applied to the input of the gate, no false pulse must appear at the output. In other words, the V_0 voltage at the logic side of the condenser must be more negative than $V_{0\,\text{sync}}$. This means that with the most positive value of $V_{1\,\text{sync}}$, the reverse current through the synchronizing diode must be sufficiently low that the voltage on the condenser remains more negative than $V_{0\,\text{sync}}$. Therefore

$$1.05i_bR_{1\,\max}=V_{0\,\text{sync}}-V_{0\,\text{in}} \tag{18}$$

In this equation, the reverse diode current is taken at a voltage $V_{1\,\text{sync}}-V_{0\,\text{sync}}$. It is apparent that the limits of excursion of the synchronizing voltage must be very closely controlled. This may be accomplished by generating a larger synchronizing pulse than is necessary and clipping the top and bottom of the wave.

To feed R_1, logic of the type discussed in this article may be used. This gate must be fed either directly or through several stages of logic by an AND gate or dummy AND gate; an AND gate resistor is necessary to supply current to charge the condenser. In case an AND gate directly feeds R_1, the following criterion may be used in determining the minimum AND gate resistor: during the initial period of charging the capacitor *C*, the voltage at the output of the AND gate must be at least as positive as $V_{1\,\text{in}}$. (This criterion may be too conservative if the choice of R_1 was dictated by the reverse current through the synchronizing diode and not by the time constant, since in this case, the initial charging current can be lower.) Note that a type of logic where V_1 is active must be used if a negative pulse is the desired output; a positive input is necessary to charge the condenser, if a rapid discharge of the condenser is to deliver a negative pulse. This means that AND gate resistors are connected to V_+ OR gate resistor to V_-.

Transmission-type gates are particularly useful where good, preferably excellent, diodes are available. Use of transmission-type gates is not recommended with diodes of low back resistance or in applications other than those involving only AND gates. Transmission-type gates are used for two purposes in this special-purpose computer:

1. Triggering a pulse amplifier from the output of a high-impedance flip-flop.

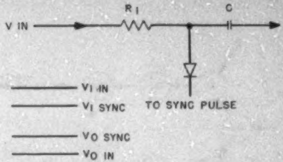


Fig. 19. Use of condenser coupling from pulse amplifiers to logic

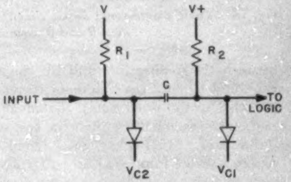


Fig. 20. Transmission-type gate and relative levels of input and synchronizing voltages

2. Triggering one of 2^n pulse amplifiers, the binary number of the desired pulse amplifier being stored on a set of flip-flops.

The method for designing transmission-type gates is much less simple and rigorous than that for designing ordinary gates. Each transmission-type gate problem must be handled somewhat differently: in some cases time is critical, in others the pulse magnitude must be carefully controlled. Generally, good diodes (especially good synchronizing diodes) are much more necessary than in ordinary gates. When a good application can be found for them, especially when good diodes are available, considerable current may be saved. It is suggested that in designing these gates, a conservative design based generally on the foregoing method be used, and that the circuit then be tested carefully.

Problem of Condenser-Coupling Inputs to Logic

Some of the disadvantages of condenser-coupling input pulse amplifiers to the logic will now be discussed. It is assumed that any pulse amplifier used for driving some logical circuit has a high impedance in one of its states: this is generally true of most circuits except cathode followers; even with cathode followers a heavy price must be paid in steady-state power drain and dissipation if the output impedance is low in both conditions.

The type of circuit to be used is shown

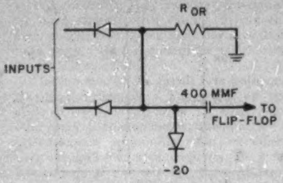


Fig. 21. Output OR gate, calculation of R_{OR}

in Fig. 20. A clamping diode on the pulse amplifier side of the condenser is desirable since it will limit the magnitude of the transmitted pulse, and hence the peak drain on the pulse amplifier.

R_2 is determined by the requirements of the logic circuits fed by the particular pulse amplifier. It is the same as the resistor required if the logic were fed from a pulse amplifier that is pulsed on rather than off. R_1 is necessary to provide d-c restoration.

The average charge flowing into the condenser is zero. During pulsing, the minimum charge flowing out of the condenser is $(V_+ - V_0)T_1/R_2$, where T_1 is the the pulsing time. During the recovery time, the maximum charge flowing back into the condenser is $(V - V_{c2})T_2/R_1$, where T_2 is the recovery time. The condenser must be large enough to pass a reasonably flat-bottomed pulse; therefore the charge on the condenser does not change appreciably during pulsing, remaining approximately $(V_{c2} - V_{c1})C$

$$\frac{(V_+ - V_0)T_1}{R_2} = \frac{(V - V_{c2})T_2}{R_1}$$

Therefore

$$R_1 = \frac{(V - V_{c2})T_2R_2}{(V_+ - V_0)T_1}$$

A clamping diode is desirable on the logic side of the condenser; if none is used a much larger voltage pulse on the amplifier side of the condenser is necessary to bring the input voltage to the logic to V_0. The necessary voltage swing is $V_{c1} - V_0$. If no clamping voltage were used, the logic side of the condenser might go well above V_{c1} to a voltage V_2; in that case, the input pulse would have to be $V_2 - V_0$ volts to activate the logic circuits. V_{c1} should be as positive as the most positive $V_{1\ \mathrm{in}}$ to any gate.

If a very large condenser is used, the ratio of T_1/T_2 can be taken as the long-time peak duty cycle; otherwise, the short-time peak duty cycle must be used. To obtain an order of magnitude, suppose $T_1 = 5$ microseconds, $T_2 = 11$ microseconds, $V_+ = 0 = V$, $V_0 = -30$, $V_{c2} = -15$; then by the time tolerances are added to the circuit $R_1 \leftrightarrows R_2$. Since only the power used to pulse R_2 can be useful to the logic, one-half the pulse power (that part going to R_1) is useless. Thus, capacitor coupling has introduced a 100% increase in peak pulse power over a pulsed-on direct-coupled pulse amplifier, and a 300% increase in peak pulse power over a pulsed-off pulse amplifier.

Supply Voltages

The problem of choosing optimum supply voltages is not an easy one. It depends on the ease of generating auxiliary voltages, the zener voltages of transistors being used, the voltage amplitude of output pulses, and the cost of power of different voltages In a system involving relatively few really long logical chains, a voltage spread of about eight times the required output swing did not anywhere introduce excessive (more than 3 to 1) attenuation from stage to stage. The greater the voltage spread, the less the attenuation from stage to stage of logic.

Low-voltage power supplies tend to be more attractive with: 1. low forward-impedance diodes, 2. few stages of logic, and 3. high-current transistor switches.

Conclusions

The advantage of multistage diode logic circuits is low cost, which is important for a computer with a complex logical system. Transistor switches ard ideal feeds for multistage circuits. In general, it is felt that multistage diode logic will have many applications in future computers.

Appendix. Calculations; Design of Transmission-Type Gates

Some Sample Calculations

In the following some typical calculations will be performed. No single chain is calculated, but some typical gates with specified output requirements are designed.

OUTPUT OR GATE FEEDING FLIP-FLOP

The requirements for the output OR gate, Fig. 21, feeding flip-flops are as follows: This gate cannot be pulsed twice in less than 36 microseconds. Most negative-input voltage is −30 volts. Use equation 19 with the modification, since in a few microseconds this condenser will be charged. $V_{\mathrm{pulse\ max}} = -20-(-30)+0$; $C = 400$ micromicrofarads.

$R_{\mathrm{OR\ max}}$

$$= \frac{36\times10^{-6}}{1.05\times400\times10^{-12}\ \ln\left(1+\frac{10}{0-(-20)}\right)} = 211K$$

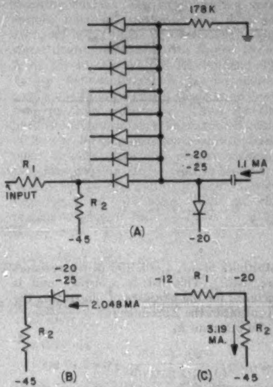


Fig. 22. Primary input-to-output OR gate

A—Entire diagram
B—Calculation of R_2
C—Calculation of R_1

In the interest of uniformity, 178K is used since, in the case of output OR gates, a certain amount of standardization may be obtained at the expense of relatively little pulse power.

DIRECT INPUT TO OUTPUT OR GATE

The following applies to a direct input to an output OR gate; see Fig. 22(A)

$$I_{\mathrm{min}} \text{ to OR gate} = 1.1 \text{ ma} + (n-1)i_b + \frac{V_+ - V_{0\ \mathrm{out}}}{(0.95)(R_{\mathrm{OR}})}$$

$$I_{\mathrm{min}} = 1.1+0.7+0.1+\frac{-(-25)}{(0.95)(178\times103)} = 2.048 \text{ ma}$$

(i_b for the clamping diode = 0.1 ma,
i_{out} to flip-flop = 1.1 ma)
V_d for 2.048 ma = 0.65 volts

The values of forward voltage drops for diodes are chosen arbitrarily for this discussion; in practice it is necessary to use a "worst" curve for the type of diode used. Therefore

$$V_{0\ \mathrm{in}} = -25-0.65 = -25.65 \text{ volts} \quad (11A)$$

$$R_{2\ \mathrm{max}} = \frac{-25.65-(-45)}{(1.05)(2.048\times10^{-3})} = 8.99K \quad (12A)$$

Use standard value 8.25K. A list of standard resistors follows:

1000
1100
1210
1330
1470
1620
1780
1960
2150
2370
2610
2870
3160
3480
3830
4220
4640
5110
5620
6190
6810
7500
8250
9090

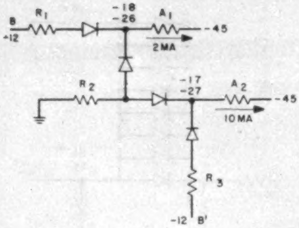


Fig. 23 (left). OR gate feeding two AND gates; both cannot be pulsed simultaneously

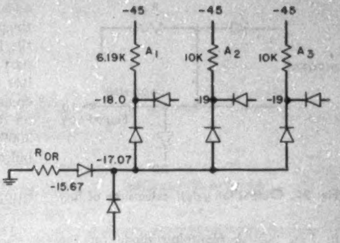


Fig. 24 (right). OR gate feeding AND gate, feeding three AND gates

Multiples of 0.1, 1, 10, 100 of the resistors are used. (This list is arbitrary and is included for the purposes of illustrating the problems in the Appendix.)

To calculate R_1

$$I_{\text{out min}} = \frac{-20-(-45)}{(0.95)(8.25\times 10^{3})} + 0 = 3.19 \text{ ma} \qquad \textbf{(13A)}$$

$$R_{1\,\max} = \frac{-12-(-20)}{(1.05)(3.19\times 10^{-3})} = \frac{8\times 10^{3}}{(1.05)(3.19)} = 2.39K$$

Use standard value 2.37*K*.

AND GATE FEEDING TWO AND GATES IN PARALLEL

Both AND gates in parallel cannot be pulsed simultaneously, see Fig. 23: R_2 must be designed so that it can supply the necessary current to A_2. R_1 can be designed so that it supplies current to A_1 when R_2 is supplying current to A_2, in case B' is pulsed.

V_d for 10 ma = 1.4 volts

$V_{1\,\text{in}} = -15.6$ volts

$$R_{2\,\max} = \frac{15.6}{(1.05)(10\times 10^{-3})} = 1.48K;\ \text{use } 1.47K$$

R_1 must be designed so that V_1 on gate A_1 when the *B* pulse amplifier is conducting is −15.6 volts; this will prevent R_2 from having to supply any current to A_1 under this condition.

V_d for 2 ma = 0.65 volts

Therefore

$V_{1\,\text{in}} = -14.95$ volts

The required current to *A*1 is somewhat higher than 2 ma because it must be supplied at −15.6 volts instead of −18 volts. Using a value of 2.2 ma for this current will more than compensate for the higher input voltage. Therefore

$$R_{1\,\max} = \frac{-12-(-14.95)}{(1.05)(2.2)} = 1.27K;$$

use 1.21*K*

CHAIN OF GATES

The following applies to a chain consisting of an OR gate feeding an AND gate which feeds three AND gates; see Fig. 24. Maximum drain from R_{OR} will occur in the case that the OR gate is the only passive input to any of the four AND gates; and in case the diode to *A*1 has a maximum forward resistance and the diodes to *A*2 and *A*3 have a minimum or zero forward resistance. This imposes the most severe current and voltage conditions on the output of the OR gate.

The maximum current to gate

$$A1 = i_b + \frac{-18-(-45)}{(0.95)(6.19\times 10^{3})} = 4.69 \text{ ma}$$

For this current, $V_d = 0.93$ volts. Therefore, the voltage applied to gate *A*1 is $-18-(-0.93) = -17.07$ volts. The current fed to gates *A*2 and *A*3 must be calculated on the basis that −17.07 volts is directly applied to the resistors of these gates, i.e., that the forward drop through the diodes to these gates is zero. Therefore

$I_{\max}$ to gates *A*1, *A*2 and *A*3

$$= 4.69 + i_b + \frac{-17.07-(-45)}{(0.95)(10\times 10^{3})} + 2i_b + \frac{-17.07-(-45)}{(0.95)(10\times 10^{3})} = 4.69 + 0.10 + 2.95 + 0.2 + 2.95 = 10.89 \text{ ma}$$

$I_{\max}$ from $R_{\text{OR}} = 10.89 + i_b = 10.99$ ma

V_d for 10.99 ma = 1.4 volts

Therefore

V_1 at $R_{\text{OR}} = -17.07-(-1.4) = -15.67$ volts

$$R_{\text{OR}} = \frac{0-(-15.67)}{(1.05)(10.99)} = 1.35K;\ \text{use } 1.33K \qquad \textbf{(15A)}$$

These calculations are necessarily somewhat sketchy. In some form, however, they present most of the problems encountered in engineering logic circuits by this method.

Example of Design of Transmission-Type Gates

In this Appendix, a typical application of transmission-type gates will be considered, with the actual design carried out.

Input devices are d-c amplifiers which may be considered switches limited in the positive direction to −15 volts, and in the negative direction to −32 volts. The d-c

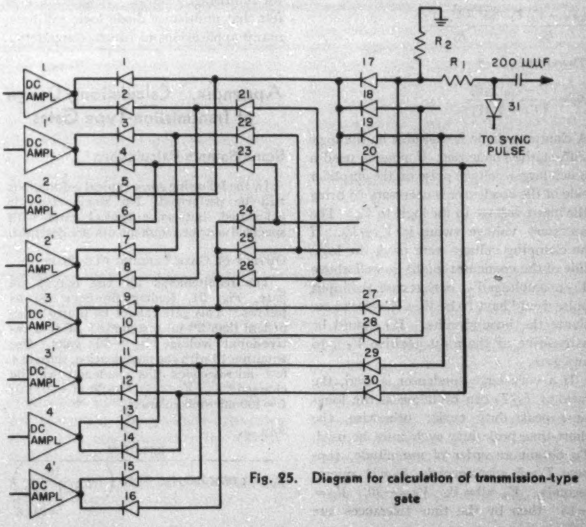


Fig. 25. Diagram for calculation of transmission-type gate

amplifiers are active when positive; when in this condition, they may be represented by open circuits with clamps to prevent their output from going more negative than −15 volts. Each d-c amplifier is attached to one side of a flip-flop.

The function of the circuit is to pulse one of 16-pulse amplifiers, depending on the state of the four flip-flops. It must pulse once every 16 microseconds; the pulse is 1.5 microseconds long. The d-c amplifiers may change 1.5 microseconds after the pulse has ended; if they change, the next pulse must represent the new condition. The pulse amplifier requires a 10-volt pulse with a 200 $\mu\mu f$ input condenser. Fig. 25 shows a typical gate and all diodes directly or indirectly associated with it.

The first step is to make an estimate of the maximum value of R_1 from time-constant considerations. Allowing three time constants:

$$3R_1C \leq (16-1.5-1.5)10^{-6}$$

$$R_1C \leq 4.33.10^{-6}$$

$$R_1 \leq 21.6K$$

Allowing tolerance, $R_1 \approx 20K$.

Now, it is necessary to check that with this value of R_1, leakage current through diode 31 when a passive signal is applied at the junction of R_1 and R_2 will not cause a false output. With −32 volts applied from the d-c amplifier, about 2 volts of forward drop through two series diodes (for example, diodes 17 and 1 in Fig. 25) must be allowed, bringing the junction of R_1 and R_2 to −30 volts. The negative synchronizing signal must be sufficiently negative so that a pulse of the correct magnitude appears at the condenser; if the current output to the pulse amplifier is 2 ma and the current to R_1 is about 1 ma, about 3 ma will flow through diode 31 in the pulsing condition. Thus, the pulsing synchronizing voltage must be about −27 volts to take care of the fact that the condenser is not charged quite to −15 originally, and to give a 10-volt pulse at the junction of R_1, and the 200 $\mu\mu f$ condenser.

When the junction of R_1 and R_2 is at −30 volts, the condenser should charge to a voltage more negative than −27 volts. Thus the leakage current through diode 31 with 12 volts across it (−15 volts from the passive sync signal to −27 volts at the junction of diode 31 and R_1), should not cause more than a 3-volt drop across R_1. If $R_1=20K$ is assumed and 0.12 ma is used as the peak value of reverse current at 12 volts, the drop across R_1 is 2.4 volts. Thus, $20K$ is satisfactory.

R_2 must be designed so that the initial voltage at the junction of R_2 and R_1 is more positive than −15 volts; otherwise, the 200 $\mu\mu f$ condenser will not charge rapidly enough. (This may not be necessary if R_1 had to be lowered to prevent false firing, since in that case, the time constant R_1C is lower than necessary.) R_2 must also supply the various reverse currents through shunting diodes. The latter are not very readily apparent, and great care must be exercised lest a diode be overlooked. The shunt diodes associated with R_2 are diodes 2, 6, 10, 14, 18, 19, 20, 28, 29, and 30. Active d-c amplifiers are low-impedance only if the external circuit tends to go more positive than −15 volts. Thus, for the purposes of determining the maximum value of R_2, the d-c amplifiers must be considered open circuits since the reverse currents through shunting diodes tend to make the voltage at the junction of R_1 and R_2 more negative than −15 volts. Thus I_m, the maximum current through R_2, is given by the equation

$$I_m = \frac{-15-(-30)}{(0.95)R_1} + 10i_b$$

$i_b = 0.15$ ma, the maximum reverse current at 15 volts.

$$I_m = 2.29 \text{ ma}$$

$$R_{2\,\max} = \frac{15}{(1.05)2.(29\times10^{-5})} = 6.24K$$

Reference

1. The Design of Switching Circuits (book), W. Keister, A. E. Ritchie, S. H. Washburn. D. Van Nostrand Company, Inc., New York, N. Y., 1951, pp. 217–26.

The Appraisal of Delays in Gate-Type Operation

IMRE MOLNAR
MEMBER AIEE

In recent years, increasing use has been made of gate-type lockout circuits in telephone switching systems. When such circuits are employed, a certain amount of delay occurs before service is given to any of the components connected to the gate; as a matter of fact one of the purposes of the gate lockout circuit is to provide a reasonably equitable distribution of delays where delays are unavoidable. To the writer's knowledge there has been no analysis published on the congestion thus introduced by gate operation, and in view of its importance in present-day circuit technique a tentative analysis is attempted in the following.

Description of Gate Operation

In switching technique frequently a call in the process of being completed must engage only temporarily one or more common-control devices. For instance, in a director system for certain types of connections each director must at one point call for the services of a translator which is common to a group of directors and can serve only one director at any one time, and the function of the director cannot proceed until it has been served by the translator. Or, as another example, in manual telephone systems every call must be answered by an operator who, upon receiving the request, establishes the connection to the called party and withdraws from the connection. In this case each call must receive service from one of a group of operators.

In such telephone systems there are two problems to be considered. First, the connection of a source to a common control must be exclusive, that is, the sources must be connected to the controls one at a time. Second, when all the common controls are occupied, the requests by other sources must be delayed until controls become available. After this waiting period, a source may be connected to a free control at random or in some prescribed order. Usually, though not always, the random connection is circuitwise the simplest, and it is satisfactory if the feature of exclusiveness only is to be safeguarded. Random service, however, is considered in most applications as undesirable, since it will not assure reasonably equitable delays among the various sources. Therefore the usual practice is to attempt to serve the sources in the same order as they have demanded service, or in some deliberate order of precedence if the sources are to be distinguished by their relative importance. The equipment required to serve calls in strict sequence of their origination is rather costly; therefore a circuit arrangement which accomplishes nearly the same result in a simpler manner is a very important one.

In a gate lockout system, calls seeking the services of common controls are held

---

Paper 55-466, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 9, 1955; made available for printing May 13, 1955.

Imre Molnar is with the Automatic Electric Company, Chicago, Ill.

in suspense as long as all controls are engaged and the previous batch of requests have not all been served. When the last one of these previous calls has been connected to a control, the gate is opened and the batch of calls waiting now is admitted. Immediately thereafter the gate is closed, and succeeding calls are required to wait until the last one of the present batch has been served. The sources admitted through the gate are served as fast as controls become available, but not in the order of their origination (otherwise it would, of course, not be a gate system). They may be served at random or (more commonly, as a matter of practical expediency) in order of precedence as determined by the relative location of each source in a circuit chain containing the totality of these sources. Many other possible arrangements are found in actual applications; however, the following analysis, based on a simple preferential chain, could serve as a pattern for the study of more complex arrangements.

From the foregoing it will be apparent that a gate system approaches, but does not attain, the objective of strict sequence service. In the progress of a call, if it is delayed at all, there are always two delay periods involved: while the call is waiting outside the gate, and while it is waiting to be connected to a control after being admitted to the gate. Of these, the first one could be extremely short in case there are no other calls waiting. As long as the first period only is considered, the calls are served in strict sequence, since it is assured that no call originating after the previous one has been admitted into the gate can obtain prior service. There is, however, no assurance that the calls inside the gate will be equitably served; only by chance would this be so.

It will also be apparent that as long as traffic is light, that is, as long as calls are being disposed of as fast as they successively come in, they are answered in strict sequence of their arrival, because as a rule there will be no call, or only one, inside the gate and perhaps one waiting outside between gate closures. This, however, is a trivial case, since under light traffic conditions calls would also be served in strict sequence if they were merely picked up at random without any gate system. On the other hand, if there is a momentary congestion so that a large number of calls collect outside the gate while the previous batch is being served, these calls are served after the next opening of the gate, without any regard to their sequence of arrival. Of course, it may take considerable time to serve all of them, and if congestion persists, a large number of calls might again be waiting for the next gate operation; therefore such congestion will be self-perpetuating, and equitable distribution will to some degree be lost. However, even under the latter condition a certain semblance of sequencing will be maintained, and better distribution obtained than if calls were served entirely at random or in a preference prescribed by their physical location in a chain.

Now, it is obvious that in a switching system, as commonly used in commercial telephone operation, there can be no prolonged periods during which all calls are consistently delayed. Such a condition would mean that more traffic was being offered than could be served by the switching devices provided; the traffic would keep backing up, and if this were to persist, say, over a whole day, the demand would never be served. Such conditions have existed and occasionally still exist in manual toll operation where the number of toll circuits is insufficient for the busy-period traffic demand. Some of the calls are then delayed until the slack period, when the number of circuits is again in excess of the traffic demand. In the meantime the strict sequence queueing is, for all practical purposes, taken care of by the operators, outside the mechanics of the switching system. Furthermore, the patience of the customer provides an effective, though very undesirable, ceiling to self-perpetuating congestion, and excessively long delays are eliminated by action of the customer in abandoning such calls.

In properly engineered systems, the numerical value of traffic (in traffic units) offered in the busy hour will be less, by a definite amount, than the number of devices provided. Therefore, during the busy hour, short spurts of congestion (during which calls are temporarily delayed) will alternate with periods during which calls may arrive singly, or there may be none at all. During such "breathing spells" the gate can clear out all calls it has stored and be ready to pick up new calls individually as they arrive. Thus, a gate system may be looked upon as strict sequence queueing between small batches of calls, and random serving within each batch. However, the probability of a large number of calls being inside the gate is much less than the probability of the same number of calls demanding service simultaneously at random when there is no gate system. When the number of calls inside the gate is small, its action is practically indistinguishable from strict sequencing; therefore it can be expected that a gate system, under usual operating conditions, will approach the results of a strict sequence system.

Indeed, experimental evidence, to be referred to later, fully confirms this expectation. In fact, in practical operation, gate systems prove to be superior to strict sequencing when (as is usually the case) the maximum number of calls being sequenced must be limited for reasons of economy. The storage means required for strict sequencing are relatively expensive, and therefore the storage capacity is generally kept as low as possible consistent with the objective. If, however, the capacity is too far limited, the purpose for which it was provided will be completely defeated. As an illustration, assume an extreme case of a sequencing circuit with means for storing two calls in their proper order of arrival. If the calls are served as fast as they come in, the correct sequence will be achieved without any queueing device. If two calls arrive while one is being processed, they will be stored and served in order of their arrival. If two more calls arrive before the first call in storage is accepted to be processed, only one of these two can occupy the vacated place in the storage circuit, and which one of the two will get it is a matter of chance. If another call arrives now, there will be two calls waiting outside the storage, and again either of the two could seize the next vacancy. Both storage positions are now occupied by random chance, therefore serving of these calls will be completely at random, in spite of the orderly sequence of their release from storage. This randomness will persist until the number of calls waiting outside of storage drops to one or zero at all times. In a mathematical sense, strict sequencing is obtainable only with an infinite storage capacity, or at least with a storage capacity equal to the number of sources originating the calls, less two (one being served, and one waiting outside the storage.) On the other hand, with the gate lockout arrangement the number of calls which can be accepted within the gate is unlimited; therefore completely randomness for an appreciable period is inconceivable.

General Considerations of Waiting Times

Delay problems in general are among the most intricate ones in traffic engineering. The solutions depend to a great extent on the distributions of the holding times, and to date for only two types of distribution have satisfactory solutions been found. The first of these deals with holding times which are exponentially distributed; there is well-established material available to analyze such problems satisfactorily.[1-5] In the second type of distribution the holding time is constant, and for this there are also computation charts available (although they are applicable only under certain restrictive conditions).[6]

The traffic analysis of gate operation is likewise a problem of delays, and it so happens that between the two previously mentioned alternatives the constant-holding-time hypothesis is the one which is more nearly correct to apply to the problems most frequently occurring in practice, since common control equipment usually has about the same holding time each time it is seized, and for the work time of an operator constancy is as good an assumption as any other. Unfortunately for the constant-holding-time case the distribution of delays has been worked out only under the assumption that the calls are served in the sequence of their origination and only up to a maximum of 20 channels, as well as with other restrictions. It would be quite useful to know the probability distribution for applications where the calls are served at random; however, no solution for this problem has yet been found.

The theory of gate operation is obviously more complex than operation either with random answering or with sequence answering, since it is a synthesis of both; therefore its exact analysis must preferably be preceded by the solution of the random answering case. The writer, being aware of the very considerable difficulties inherent in the problem, as well as of his own mathematical limitations, in the following tries only to give an outline of a theory with gross approximations. There are only two excuses offered for doing this: first, as long as there is no exact theory available or in sight, the practical engineer who must face such problems may prefer to have a reasonable approximation rather than no guidance; second, this presentation may arouse the interest of others, mathematically better equipped, to develop a satisfactory theory and reduce it to working material.

In support of this position, the following testimonial from the seventeenth-century French philosopher, René Descartes, may be quoted: "I am sure that there is no one, even among those who make its [medicine's] study a profession who does not confess that all that men know is almost nothing in comparison with what remains to be known. I judge there was no better provision against a short life or lack of experience than faithfully to communicate to the public the little which I should myself have discovered, and to beg all well-inclined persons to proceed further and then to communicate to the public all the things which they might discover, in order that the last should commence where the preceding had left off, and thus by joining together the lives and labors of many, we should collectively proceed much further than anyone in particular could succeed in doing."

For example, consider the position of the practical designer who is expected to lay out some new type of manual board. He knows that there are a certain number of calls to be answered in the busy hour, and can guess the amount of work time involved in each call. He is also told that 95 per cent (%) of the calls must be answered within 10 seconds. The first thing he wants to know is the number of operators necessary to carry this load, and for this he might use some empirical figures, or he could look up the charts prepared from the mathematical material given in reference 6.

What he has not been told is the sequence in which these calls are to be answered and if, after the board has been put into service, an undue number of calls are abandoned because of excessive answering delays (even though in the over-all picture 95% are answered within 10 seconds) the designer is in trouble. He would not be, had he arranged for strict sequence answering, except that very likely the board would be economically unattractive, as has been noted, because of the high cost of an effective sequencing arrangement.

So he considers the use of a gate lockout system, and wonders where this will lead him. Intuitively he feels that this is almost as good as strict sequence; but if he is conscientious, he will ask himself how many calls on the average will collect outside the gate, and how many will be inside; how likely is it that, say, five times the average number will wait; will a gate system prolong the average delay; and how long will the gate be held between successive operations? He must answer these questions not only for his appraisal of the efficiency of the system, but also for a number of seemingly trivial reasons, such as contact protection, voltage drop through common impedances, provision of time delays for alarms under trouble conditions, etc.

At present he does not have even approximate answers. He does not know whether at any one time there will be 1, 2, 5, or 50 calls waiting in the gate; or even, if he should know that on the average there are, say, two, whether he can reasonably expect that there ever will be 50, and how far this would throw out the sequencing. If someone can tell him, for instance, that in a given case there will be on the average between 1.5 and 2.5 calls in the gate, and that about once a week he might have more than five calls, and it should practically never happen that there are more than 10 calls, that there should not be more than 25% spread in the average delays between the first and last members of the preference chain, etc., then he has something to work with and something on which to base his decisions.

This is not an apology for the lack of mathematical sophistication in the following appraisal. There is no substitute for rigor in analysis, and there is always a danger of getting a completely misleading result by careless reasoning and half-knowledge; however, it must be recognized that there is real value in any reasonably true answer until a better one with a more solid mathematical justification can be found. The exactness of the answers given to questions which arise out of gate operation have little effect on the cost of the equipment, or the exact grade of service given to the customer, or the specifications to be met in competitive projects. Indeed, an answer which is correct to any order of magnitude helps the designer to visualize the events to take place and gives him valuable guidance in making decisions on the details of his design.

In the following it will be assumed that statistical equilibrium exists; that the traffic originates from a very large number of sources; that there is full availability from source to channels (gate and controls); that the connection from sources to channels is instantaneous; that calls are originated at random; that no one source is preferred over any of the others as long as they are waiting outside of the gate but that, once admitted to the gate, they will be served in a certain fixed order of preference assigned to each source; that calls enter the gate only when all channels are busy or when other calls are waiting. In general there will be no restriction placed on the form of the holding-time distribution, though several of the suggested answers will be tied to the case of constant holding time.

List of Symbols

s = number of sources
c = number of calls arrivals during unit time
y = traffic density in traffic units, erlangs
T = average channel holding time, unit time
N = number of channels
m = average delay of all calls, unit time

M = average delay of delayed calls, unit time
G = total gate-holding time, unit time
g = average gate-holding time, unit time
k = number of gate operations during unit time
D = number of delayed calls during unit time
h = average number of delayed calls during G
H = average number of calls entering gate
n = number of calls entering gate
b = average delay outside of gate, unit time
t = time variable, unit time
$P(>0)$ = probability for delay
$P(>t)$ = probability of delay exceeding t
$Q(>t)$ = probability of gate-holding time exceeding t
$R(>n)$ = probability of at least n calls entering gate
$P(x)$ = probability of x channels being simultaneously occupied
K = multiplier applicable to gate operation
$\epsilon = 2.718282$
S^* = summation

Proportion of Delayed Calls

As is well known, in a full availability system the total amount of delay suffered by all calls does not depend upon the order in which the calls are being served, and this will also be true if the delayed calls pass through a gate lockout process. Furthermore, it should make no difference to the undelayed calls in what sequence the delayed ones are served. Therefore the existing equations for the probability that a call will not be delayed are still applicable; and so are the probabilities that a call will be delayed, since the sum of these two probabilities is identically equal to 1. The probability for a call to be delayed is designated as $P(>0)$. For the exponential holding case

$$P(>0) = P(0)y^N/(N-y)(N-1)! \qquad (1)$$

where

$$1/P(0) = \epsilon^y + y^N/(N-y)(N-1)! - \overset{\infty}{\underset{x=N}{S^*}}\; y^x/x! \qquad (2)$$

For the constant holding-time case

$$P(>0) = 1 - P(=0)$$

where

$$\log P(=0) = -\overset{\infty}{\underset{x=1}{S^*}}\,(\epsilon^{-xy}/x)\overset{\infty}{\underset{j=xN}{S^*}}\,(xy)^j/j! \qquad (3)$$

Numerical values of the functions in this and the next section are contained in the references, for all cases of practical importance and to an adequate degree of accuracy

The probability $P(>0)$ represents the time during which it is expected that all channels will be engaged and, since calls are originated at random, it also expresses the expected proportion of delayed calls out of the total number originated. It does not, however, represent the time during which delays will actually occur, since there will be some periods when all channels are busy and no calls are waiting.

Average Delay

An examination of the equations for the average delay of all calls shows that they are independent of the manner in which delayed calls are served and therefore applicable also to the present problem. In the exponential case

$$m = TP(>0)/(N-y) \qquad (4)$$

In the constant holding-time case

$$m = T\overset{\infty}{\underset{x=1}{S^*}}\,\epsilon^{-xy}\left(\overset{\infty}{\underset{j=xN}{S^*}}\,(xN)^j/j! - (N/y)\overset{\infty}{\underset{j=xN+1}{S^*}}\,(xN)^j/j!\right) \qquad (5)$$

Numerical values of the average delays are likewise given in the references for a very wide range of conditions.

The delays m refer to all calls, whether delayed or undelayed. The average delay of those calls which are delayed is

$$M = m/P(>0) \qquad (6)$$

which is simply $T/(N-y)$ for the exponential holding times.

The delays with exponential holding times are consistently higher than they are when the channel holding times are constant, assuming the same average holding time in both cases. This also bears out intuition that with exponential distribution, holding times of lengths many times the average will occur and create excessive delays, while, when the holding times are all equal in length, delayed calls will be served in regularly occurring equal intervals. As a rule of thumb, up to about 10 channels and high occupancy (y/N between, say 0.7 and 0.95), the average delay m is about twice as long with exponentially distributed as with constant holding times. With higher occupancy the ratio also increases, while with lower occupancies and also with increasing number of channels, the average delays with both types of holding-time distribution approach each other. This is likewise in accordance with intuition, since when the occupancy is low, most calls will be served without any delay, and any differences in the delay distribution of those calls which will be delayed become less significant in the overall average as the proportion of delayed calls diminishes.

The mathematics of delays becomes very simple in the special case of a single channel. However, this case happens to have special importance, since gate lockout systems are common in single-channel operation. In this case, with exponential holding time distribution

$$P(>0) = y \qquad (1A)$$

$$P(0) = 1 - y \qquad (2A)$$

$$m = yT/(1-y) \qquad (4A)$$

$$M = T/(1-y) \qquad (6A)$$

With constant holding time

$$P(>0) = y \qquad (3A)$$

$$m = yT/2(1-y) \qquad (5A)$$

$$M = T/2(1-y) \qquad (6B)$$

As may be seen, the probabilities of delay are the same in both cases, while the average delay in the exponential case is twice as long as with a constant holding time for all levels of occupancy.

Distribution of Delays

As a practical measure the distribution of delays is of much greater importance than the average delay. It is very little consolation to know that in a given system the average delay is a comfortably small figure, if at the same time an undue number of calls are excessively delayed. A specification for a system where calls might be delayed to a varying degree should always include a provision that not more than a certain fraction of calls should be delayed over a certain length of time; and it is expedient to express this delay t in multiples of the average holding time T.

A typical example of such a specification may be given by a common control system where a large number of directors must seek the services of a common translator for every call which the directors process. Let us assume that every call occupies the translator for the same length of time, say 350 milliseconds, including access time. The question is now asked as to how many calls the translator will be able to process in the busy hour. If the translator should serve 3,600/0.35 = 10,286 calls, it would be fully occupied and delays of infinite length could be expected. The permissible number of calls, less than this amount, depends on the length of delay which may be tolerated; it may be specified, for instance, that not more than one call out of 10 shall be delayed for more than twice the holding time, that is, for 700 milliseconds. The question can be answered if the distribution function of delays for the exact process by which the translator is accessed is known, with an accuracy limited by the correctness of the assumptions on which the distribution function was based and by other general limitations on the application of the theory of probability to practical problems.

The equipment might be arranged so that calls would seize the translator at random (in which case some extremely long delays might be expected) or in strict order of their origin (however, the equipment would then be very costly) or in a gate sequence arrangement. From previous qualitative considerations it is expected that the delay distribution function for gate operation should lie between those for random and for strict sequence service respectively. However, it would be presumptuous to guess its exact position in the absence of a mathematical formulation or of sufficient statistical material. One might be tempted to interpolate about halfway between the two functions, on the reasoning that delayed calls spend about half their time outside and half inside the gate, so they would be served about half in sequence and half at random. This reasoning is, however misleading, first, because delay distribution functions are basically summations of exponential terms, and linear interpolation of such functions is completely unrealistic; and second, because intuition as well as the little statistical evidence available clearly suggest that a gate sequence is a much closer approach to strict sequence in all cases of practical interest.

Distribution functions and their numerical values are available[1–5] for the exponential distribution of holding times, both for strict sequence and random services. Unfortunately, however, for the constant holding-time case (which is the more common one in gate system practice) the values for strict sequence service only have been worked out and are available for reference.[6]

In approaching the gate distribution problem, the qualitative observations in the paragraphs to follow may be useful:

Delay distribution functions demonstrate that with random answering a larger proportion of calls have shorter delays than with strict sequence answering. This is the mathematical justification of strict sequence service, for in addition to providing greater equity, it eliminates excessively long delays. That this must be so is obvious from the fact that the average delay is the same in both cases, and therefore the reduction of long delays must be compensated by a corresponding increase of the shorter delays, with both distribution curves crossing over at some point.

From the fact just mentioned it follows, somewhat surprisingly, that if the possibility is accepted of some very long delays with random service, a greater proportion of calls can also then be served with little delay than is possible with a sequencing arrangement. This seems to have been overlooked until now by designers, and opens up the possibility of new and interesting sequencing combinations.

As already noted, with a low occupancy of the channels a gate lockout arrangement should be nearly as effective as strict sequencing, while with increasing occupancy more randomness is introduced. As a matter of fact, the available distribution functions demonstrate that at low occupancies there is little difference of practical consequence even between strict sequence and random service.

More specifically, consider the exponential holding-time case, for which exact mathematical comparisons can be made, and three probability levels: 0.1, 0.01, and 0.001, which should include the total range of practical significance. At the outset it may be well to remember that the probability of all channels being occupied, that is, the probability that calls will be delayed, is independent of the process by which delayed calls are disposed of in switching systems used at present. This probability is shown in Table I, which indicates that with low occupancies, say 0.3 or 0.4, and with 10 channels or more, the question of random or queued service is completely irrelevant. With very low occupancies, say 0.1 or 0.2, this is true even with four channels or more. The values are even more favorable in the constant holding-time case.

The next step is to investigate the maximum occupancies of the channels permitted if the delays with random service are not to exceed the delays with strict sequence service by, say, 50%. It will be found that at the 0.1 probability level the occupancy will be over 0.9 for any number of channels. At the 0.01 probability level the occupancy will be about 0.8 for any number of channels, and will rise to about 0.9 for 20 channels. At the 0.001 probability level the occupancy will be about 0.5, rising to 0.6 for 10 channels, and to 0.75 for 20 channels. It may be further observed that up to 0.9 channel occupancy, and any number of channels, the delay with random service at the 0.001 probability level will be less than about twice the delay with strict sequence service.

Since there is no analytical expression for the distribution of delays with gate-type operation, the next best procedure is to apply multipliers K to the numerical values of delays derived for strict sequence service, for the same number of channels, for the same amount of traffic to be carried, and at the same probability level $P(>t)$. As has just been seen, the delays with random service will not exceed the delays with strict sequence service by more than about 50% for the major portion of cases of practical interest, and in all cases of practical interest, will not exceed them by more than 100%. As it was also previously concluded that the distribution function for gate-type service should be closer to strict sequence service than halfway between strict sequence and random distributions, it seems reasonable to assume that for engineering purposes the multiplier K should vary between 1 and 1.3, dependent on the three parameters mentioned.

It is assumed in the foregoing that delays with gate-type operation will be longer than with strict sequence queueing. Actually, as was pointed out in the beginning of the discussion, with certain values of the parameters the delays with gate sequence will be shorter; therefore disregarding the possibility of multipliers less than one is conservative engineering practice.

The considerations were based on exponential holding-time distribution. It seems reasonable to assume that with constant holding time the spread will be appreciably less. Therefore the application of the same values of K to all types of holding-time distributions is further on the conservative side.

Whether or not the specified range for the multiplier K is reasonable can be decided, in the absence of a solid analytical foundation, only from statistical observations. To undertake a statistical survey of this problem is a formidable task, considering the wide range of two parameters involved (number of channels and average occupancy) and the huge sample size required for each value of the parameters to obtain meaningful results at the probability level 0.001 for $P(>t)$. Fortunately, there are data available from one experimental investigation.[8] Among

Table I. Probability of Delayed Calls

Probability	Number of Channels	Average Occupancy
0.1	1	0.1
	2	0.25
	10	0.6
	20	0.7
0.01	2	0.075
	4	0.2
	10	0.4
	20	0.55
0.001	4	0.1
	10	0.3
	20	0.45

other aspects, the effect of gating was studied on an artificial traffic of 1,000 calls generated by random numbers and served by two channels. The average occupancy y/N was 0.9. Such high occupancy with only two channels represents, of course, a much more severe test than can be commonly found in practice. Therefore, subject to the usual statistical limitations on such modest sample sizes, the results are conservative compared with practical operations. The result of this experiment showed that with gate sequence service, and at probability levels below 0.1, delays were about 30% longer than when the same test calls were served in strict sequence.

A few things to keep in mind in considering the results of this experiment are: 1. The number of samples at low probability levels was very small (0.01 is represented by ten samples, 0.001 by only one). 2. In the probability range of interest in this connection, all the delays obtained in the test were consistently lower than the theoretical values. The discrepancy was less with random service (about 15–20%) and much greater with strict sequence service (30–50%). This suggests that the gate sequence delays may tend to be closer to strict sequence delays then the test indicates. 3. All delays obtained in the test for gate sequence, in the probability range under discussion, were less then the theoretical delays for strict sequence service.

The discussions in this section may be summed up by suggesting the multipliers K shown in Table II for engineering purposes.

In conclusion, the previous example may be continued of a translator to serve calls of 350 milliseconds' holding time, with not more than one in 10 delayed over 700 milliseconds. From the distribution curves contained in reference 6 it is seen that with strict sequence service an occupancy of 0.58 will satisfy the specification. At this occupancy $K=1$, so the same value is also applicable for gate sequence service, and will be obtained by 6,000 hourly calls of 0.35 second each. From these charts it is also seen that one call out of 100 will be delayed over 1.54 seconds, and one call out of 1,000 over 2.6 seconds. In the last value the multiplier $K=1.1$ has already been applied in accordance with Table II; without it the delay would have been over 2.38 seconds. Had establishing numerical values for "K" been greatly in error, assuming, for instance, 1.3 the correct value instead of 1.1, the delay for one call in 1,000 would increase to over 3.1 seconds. While the discrepancy in per cent is substantial, it is of little practical significance whether the delay is 2.6 or 3.1 seconds for one call in 1,000.

However, if the specification had called for one call in 1,000 delayed for over 2.6 seconds, instead of the original one call in 10 for over 0.7 second, and had there been an error with regard to the multiplier, in that it should have been 1.3 instead of 1.1, then the permissible occupancy would have reduced from 0.58 to 0.53, and the permissible number of calls from 6,000 to 5,450.

Table II. Multipliers

Probability P(>t)	0.1		0.01		0.001	
Occupancy y/N	0.5	0.9	0.5	0.9	0.5	0.9
K for one or two channels	1.0	1.1	1.0	1.2	1.1	1.3
K for over 10 channels	1.0	1.0	1.0	1.1	1.0	1.2

Preference Service

It has been specified that, once inside the gate, calls are served according to the preference assigned to the sources by their physical arrangement in a lockout chain. However, this is equivalent to random service within the gate, because the association of the preference number with the time of origination of a call is completely at random. Therefore, if these calls were to be subsequently served in a sequence of their precedence, the over-all gain or loss of the group as a whole compared with strict sequence service would be the same as if they had been served entirely at random.

To illustrate, the following two experiments may be considered. Of two decks of cards, each consisting of, say, 26 cards, the cards in the first deck are marked on one side with letters A, B, ... through Z, where the letter designates the rank of the card, A being the highest ranking. After thoroughly shuffling them a few, say five, are picked at random and each one is marked on the other side by the numbers 1, 2, ...5 according to the order in which they are drawn. Now the cards are dealt one at a time in the order of their rank, corresponding to their letter designation, and the number is marked down of places gained or lost by each card as against dealing them out in order 1, 2, ...5.

For the second experiment the 26 cards of the other deck are completely blank on both sides. Again, five cards are picked out which could be from the top or bottom or from any place in the deck since, being blank, the cards are indistinguishable. As the cards are drawn, again they are marked 1, 2,...5. The cards are not shuffled before being drawn out. They are now shuffled and dealt singly in any order and the number assigned to them is not looked at. To make this experiment identical with the previous one, each card before being dealt may have a different letter assigned to it in any arbitrary way and then dealt according to the rank of the letter assigned to it. Finally, again the gain or loss for each card is noted as compared with its respective number, 1, 2...5.

If the conditions of both experiments are checked, it will be noticed that they are completely equivalent, since in the first one the five letters were pulled entirely at random out of a possible 26, while in the second, five letters out of a possible 26 were assigned entirely at random to the selected five blank cards. In the first experiment the cards were dealt according to their preference in rank. In the second, the cards were thus dealt but since the letters were assigned at random, they might just as well have been dealt at random. It may therefore be expected that after many repeated trials the distribution of gains and losses will be the same for both sets.

These relations can also be expressed in mathematical language. For this purpose it is assumed that an index 1, 2, through s is assigned to each source according to the preference by which it gains access to a channel, 1 being the highest ranking source. A call is originated by source of index i during an all-channel-busy condition and with $x-1$ sources already waiting for admission into the gate. Once this call is inside the gate, there are altogether n sources waiting for service, and it can be expected that $n(i-1)/s$ calls of these sources are ranking higher than i. Now, obviously, the number of calls in the gate ranking higher than i must be an integer number, while $n(i-1)/s$ will be a rational number but only occasionally an integer. The integer number of calls in the gate which rank higher than the call of our test source i, and which therefore will be served prior to i, is $(n(i-1)/s)-r_i$ where $0 \leq r_i < 1$ and represents the difference between $n(i-1)/s$ and the next smaller integer. Our test source will gain $x-1+r_i-n(i-1)/s$ places in its delay over what it would have been entitled to, had the calls been served in sequence of their arrivals (negative gains representing, of course, a loss in place). The test source was chosen at random with a probability of $1/s$, and the expected gain can be obtained by summing up over all values of i, from 1 to s, as follows:

$$x-1-\overset{s}{\underset{i=1}{S^*}}\, n(i-1)/s^2+\overset{s}{\underset{i=1}{S^*}}\, r_i/s$$
$$=x-1-n/2+n/2s+\overset{s}{\underset{i=1}{S^*}}\, r_i/s$$

The last sum in this equation can be computed by considering consecutive values of r_i. Now, $n[i-(i-1)]/s=n/s$, which is equal to $r_i-(r_i-1)$, except in the trivial case when $n=s$. A sufficient number of terms, say k, are being collected such that $kn/s=1$; that is, $k=s/n$. In considering the complete set of s sources this will be repeated cyclically altogether $s/k=n$ times. The sum $\overset{s}{\underset{i=1}{S^*}}\, r_i$ will consist of n equal sums of an arithmetic progression, each consisting of $k-1$ terms: $\overset{k-1}{\underset{j=1}{S^*}} nj/s=nk(k-1)/2s=(s-n)/2n$. Therefore the sum $\overset{s}{\underset{i=1}{S^*}} r_i/s={}^1/_2-n/2s$. Substituting this the following is obtained

$$\text{expected gain}=x-(1+n)/2 \qquad (7)$$

Had the test call been served at random, without regard to the rank of the source, there would have been an equal probability $1/n$ that 0 or 1 or... $n-1$ calls would be served prior to our test call. The expected gain in this case would be

$$x-1-\overset{n-1}{\underset{z=0}{S^*}}\, z/n=x-n/2-{}^1/_2$$

Thus the equivalence of both distributions is proved.

It should be emphasized that while there is no difference between these two distributions as far as the proportion of all calls delayed for t or longer is concerned, it certainly makes some difference to the expectation of individual sources; obviously, higher ranking calls will be subjected to shorter average delays than those behind them in the preference chain.

Gate-Holding Time

The length of time $P(>0)$ has already been considered, during which all channels are engaged and a delay can be expected. The gate holding time G cannot entirely cover this period, because at least part of this time no new calls will arrive and no other calls will be waiting. The gate will be held during the time that at least one call is waiting. The unconditional probability for this is the product of the probability $P(>0)$ that all channels are busy and the conditional probability that one or more calls are waiting. The first probability can be obtained from equations 1 or 3 or from charts. The conditional probability that no calls are waiting when all channels are busy is $1-(y/N)$, therefore the probability that at least one call is waiting is y/N. The total gate occupancy during unit time will be

$$G=P(>0)y/N \qquad (8)$$

Average Number of Delayed Calls

The total amount of delay in unit time is mc, while the total gate occupancy from equation 8 is G. Since the gate is closed whenever calls are delayed, the average number of calls waiting during G is mc/G, or

$$h=Nmc/yP(>0)=NM/T \qquad (9)$$

As h represents the average number of calls waiting at any time, it will also be the average number of calls waiting outside the gate at the time when the last call was cleared out of the gate and the gate reopened to admit a new batch of waiting calls. Therefore, the average number of calls entering the gate is

$$H=h=NM/T \qquad (10)$$

Number of Gate Operations

The total number of delayed calls is

$$D=cP(>0) \qquad (11)$$

while in the average $H=NM/T$ calls are admitted to the gate. Therefore the number of gate operations will be

$$k=D/H=yP(>0)/NM \qquad (12)$$

The total occupancy of the gate is G, and there are k separate closures. Therefore the mean duration of each occupancy of the gate will be $g=G/k$, and by substitution of equations 8 and 12

$$g=M \qquad (13)$$

This result could also have been obtained by intuition, since the average delay of a delayed call is composed of a delay outside the gate until the gate is reopened, plus a delay inside the gate until served, which on the average will completely cover one full gate occupancy.

Average Delays Outside and Inside the Gate

The average holding time of the gate being g, the conditional probability for a certain call to originate during a small interval dt at a time t after the preceding closure of the gate, if it is known that it originates during g, is dt/g. This call would then have to wait $g-t$ until the next opening of the gate. The expected delay outside the gate would be

$$b=\int_{t=0}^{g}(g-t)dt/g=g/2 \qquad (14)$$

thus from equation 13

$$b=M/2 \qquad (14A)$$

This means that a call spends, on the average, one half of its expected delay M (if it is known that it will be delayed) outside the gate; and therefore, obviously, the other half of its expected delay inside the gate.

Distribution of Gate-Holding Times

Now, to prove the probability $Q(>t)$ that the gate is closed for a duration of least t, it may be assumed for the moment that some arrangement is made whereby the calls inside the gate will be served in the order of their time of arrival. Of course, this would not be sensible since obviously the gate will serve no useful purpose. Nevertheless, with such an arrangement the calls are served in strict sequence and the probability of their delays $P(>t)$ can be readily obtained from the available charts for strict sequence serving. The probability of a delay of at least t on any call which is known to be delayed will be $P(>t)/P(>0)$. This probability applies to any call as long as it arrived at a time when all channels were engaged, even with the special gate circuit (which is now completely ineffective on the sequence of serving the calls). It is therefore equally applicable to a call which happens to be the first one to arrive after the gate was closed. With the special arrangement this call will be the first one to be served after the next opening of the gate; therefore it will be delayed for the duration of the gate-holding time.

The gate-holding time obviously, depends only on the number of calls to be served within the gate, the number of channels available, and the holding time of the channels, and not on the sequence of serving the individual calls. The gate-holding times will therefore be the same with the special sequencing arrangement and with ordinary gates, if the number of calls admitted to the gate is the same in both cases. On the other hand, the number of calls admitted depends on the preceding gate-holding time. Since calls are originated at random, the number of calls seeking admission depends on the calling rate of the sources as well as the duration of the gate holding time. Because the sources originating the calls have no knowledge of the state of congestion or of the method of sequencing the calls, the calling rate will be the same for all methods of serving the calls. Furthermore, since the length of the gate-holding time would not depend on the method of sequencing unless the method should affect the number of calls collecting during the gate-holding time, the equality of gate-holding times with both methods of service has been demonstrated.

The special arrangement of serving the calls within the gate in sequence of their arrival may now be dispensed with. Then the probability $Q(>t)$ that the gate will be held for at least t is equal to the probability that delayed calls will be delayed at least t if they are served in the sequence of their arrival. If $P'(>t)$ is the distribution for strict sequence service, then

$$Q(>t)=P'(>t)/P(>0) \tag{15}$$

Note that $Q(>0)=1$, and that the average gate-holding time g is M by using equation 6, and in conformance with equation 13. For exponential holding-time distribution

$$Q(>t)=\epsilon^{(y-N)t/T} \tag{15A}$$

In the constant holding-time case the expression for $Q(>t)$ is more complex and should preferably be obtained from the suggested charts.

Delays Outside the Gate

To derive an expression for the probability that a delayed call must wait longer than t to be admitted into the gate, consider a gate-holding time i. Calls being originated at random, the probability that a call will have to wait for more than t for the gate to open is $(i-T)/i$. If the probability of a gate-holding time between $i+di$ and i is $Q(i)di$ the total probability that a delayed call will be delayed outside the gate for longer than t is

$$\int_{\infty}^{i=t} Q(i)di-t\int_{\infty}^{i=t} Q(i)di/i$$
$$=Q(>t)-t\int_{\infty}^{i=t}Q(i)di/i$$

The evaluation of the integral causes difficulties. For instance, in the exponential holding-time case, which appears to be the simpler one, $Q(t)$ can be obtained from equation 15(A). The integral is solvable by a power series expansion, plus a term log i. Whether the series converges at the limits of integration was not investigated.

The solution of the foregoing equation, as well as the corresponding one for the distribution of delays inside the gate, would have considerable importance, since the combination of both distribution functions would provide $P(>t)$, and thus the complete solution of the gate problem, instead of the intuitive approach outlined in the section "Distribution of Delays."

An estimate of the probability distribution of delays outside the gate may be obtained as follows. As noted, on the average, a delayed call spends one half of its expected delay outside and one half inside the gate. Therefore, as an engineering approach, it may be considered that the probability of a delayed call being delayed outside the gate for more than t is the same as the probability for gate holding times over $2t$, or $Q(>2t)$.

Distribution of Calls Entering the Gate

If there are n calls admitted into the gate, these calls must be served during the subsequent gate holding time. Therefore, the probability $R(n)$ for the number of calls entering the gate should be the same as the probability $Q(t)$ for a gate holding time t, if the latter is computed so that on the average n calls will be served in that time. The meaning of $Q(t)$, of course, is that the probability for a gate holding time between t and $t+dt$ during the time that all channels are engaged is $Q(t)dt$; its value can be computed from the relation

$$Q(t)=-dQ(>t)/dt$$

$R(n)$ represents the proportion of gate operations when exactly n calls are admitted into the gate. It can be computed as follows. Consider first gate holding times of a length between t and $t+dt$ and compute the probability that exactly n calls will be served in the gate. This probability should then be multiplied by the probability $Q(t)dt$ that the holding time will be of such length, and the products integrated for all possible values of t.

With exponential channel holding time distribution the probability of a call terminating during any interval dt at any time is equal to Ndt/T, irrespective of the length of time the calls might have been in progress. On the average, we may expect $x=Nt/T$ calls to terminate during an interval of length t. The probability that during a gate holding time t exactly n calls will terminate is $x^n\epsilon^{-x}/n!$ Therefore, regardless of the length of gate holding time, the total probability that n calls will be admitted to the gate is

$$R(n)=(T/N)\int_{x=0}^{\infty} Q(t)x^n\epsilon^{-x}dx/n!$$

From the theorem of the mean for integrals, there will be a value t' such that

$$R(n)=Q(t')T/N\int_{x=0}^{\infty}x^n\epsilon^{-x}dx/n!$$
$$=Q(t')T/N$$

The interpretation of t' can be obtained by the following consideration. Because the number of calls in the gate must be an integral number, the probability on the right-hand side must change in discrete steps as n changes. Since T/N represents the average interval between channel releases, $Q(t')T/N$ is the probability of the gate holding time for n calls. $Q(t)$ being a monotonous function, t' must have a unique value for each n. Now, it is known that if $t=nT/N$, the expected number of calls in the gate will be n. The foregoing equation should certainly still be satisfied if this constant value is assumed for all gate holding times during which n calls are served.

To obtain the probability of at least n calls admitted to the gate $R(>n)$ we must add $R(n)$ for n, $n+1,\ldots.$ Obviously, $R(>n)$ will then be $Q(>t)$, where t'' is some value of t near t'. The distribution function $Q(>t)$ being continuous and regular; in fact, it is $\epsilon^{(y-N)/T}t$ must be established in such a way that the infinite sum of $Q(t')T/N$ will be $Q(>t'')$. Now, $R(n)=R(>n)-R(>n+1)=Q(>t'')-Q(>t^*)$ where $t^*=t''+T/N''$. Substituting $\epsilon^{(y-N)t/T}$ for $Q(>t)$

$$R(n)=\epsilon^{(y-N)t''/T}(1-\epsilon^{-(1-a)})$$

if $y/N=a$, the average channel occupancy.

On the other hand

$$R(n)=Q(t')T/N$$

and

$$Q(t')=-dQ(>t)/dt \text{ at } t=t'$$

Thus

$$R(n)=(1-a)\epsilon^{(y-N)t'/T}$$

and, equating both expressions for $R(n)$

$$\epsilon^{(y-N)(t''-t')/T}=(1-a)/(1-\epsilon^{-(1-a)})$$

$$T''-t'=[\log(1-a)-\log(1-\epsilon^{-(1-a)})]$$
$$T/(y-N)$$

natural logarithm understood throughout t' is the mean value of t in the interval of width T/N beginning with t'', and is nT/N. For high-channel occupancies, which are usually of interest in such problems, $\epsilon^{-(1-a)}$ is approximately a, thus $t''-t'$ approximately 0. Therefore for practical purposes

$$R(>n)=Q(>t), \text{ if } t=nT/N \tag{16}$$

In the constant channel holding-time case it is known, by hypothesis, that during a gate holding time T, equal to the channel holding times, N calls must terminate; therefore during an interval t which is a multiple of the channel holding time tN/T calls must terminate. The probability of n calls admitted into the gate will be the same as the probability of a gate holding time t, where $t=nT/N$, if

n is a multiple of the number of channels N.

If the number of calls admitted is less than the number of channels, the gate holding time will be less than T and the gate will be reopened as soon as n of the channels become free. How long this will take depends on the exact instant these channels became occupied during the preceding gate occupancy, because each channel will be held for the exact length T. However, since N channels must become free during T, it can be expected that the n calls will be served in $t=nT/N$, which is then the expected gate holding time when n calls have been admitted. Contrariwise, the probability of admission of n calls less than N will be the same as the probability of gate holding time $t=nT/N$ when t is less than T.

Similar considerations apply when the number of admitted calls falls between two multiples of N. In general, it can be estimated that the probability of more than n calls being admitted to the gate will be in accordance with equation 16.

The average number of calls entering the gate will be

$$\overset{\infty}{\underset{n=0}{S^*}} nR(n)=(N/T)\int_{t=0}^{\infty} tQ(t)dt=MN/T$$

in accordance with equation 10.

Average Delay of Individual Sources

While in general the service inside the gate can be considered as being at random as far as the over-all probability distribution of delays is concerned, this will not apply to the individual sources if they are arranged in a certain order of preference. It was shown that the average delay outside the gate for any source is $g/2-M/2$. Once inside the gate, the highest ranking source will only wait for any one channel to become free, which on the average will be in T/N, while the lowest ranking source will be served in time g after the closing of the gate.

This might appear to be a contradiction of the previous conclusion that the average delay inside the gate is $g/2$, since from what has just been said the average delay will obviously be greater than $g/2$. If the average delay of these calls happened to be M, their average delay outside the gate must have been less than $g/2$, with this as well as the preceding gate holding times $g=M$. Now, equation 14 is undoubtedly correct with random incidence of calls, but since the average delay inside the gate was greater than $g/2$, and provided that two consecutive gate holding times were g each, the average delay of these calls must have been greater than $g=M$. The contradiction could exist only if the gate holding times were persistently equal, which would mean a persistently uniform flow of traffic. During such a state the channels would persistently be fully occupied and (dependent on the initial condition of the state) any value of average delay higher than the gate holding time would be obtainable. However, the existence of such a persistent state is in contradiction to the hypothesis of the delay problem, according to which varying amounts of traffic are being offered to the system, and therefore varying lengths of gate occupancies will be obtained. These occupancies will adjust themselves so that, on the average, the relation expressed in equation 13 is satisfied.

Disregarding the delay T/N for the highest ranking source, the expected spread of the average delay between highest and lowest ranking sources will be from $g/2$ to $3g/2$. The expected delay for the source of rank i, if there are altogether s sources, is

$$g/2+g(i-1)/(s-1) \tag{17}$$

In conclusion, it should be repeated that many details of the foregoing analysis are indefensible as to mathematical rigor. No claim can be made that the gate problem is solved as a mathematical subject; however, enough material is given to proceed with engineering design with a reasonable assurance of safety even though the exact factor of safety present cannot be determined.

Examples

The following two examples will illustrate the application of the material to typical problems in telephone engineering. The numerical values of the probability functions used in these examples will be obtained from the reference material,[8] which the reader should consult if he wishes to verify the manner in which these values were obtained.

EXAMPLE 1

To continue the example of the common translator with a holding time of 350 milliseconds, this time it will be assumed that 7,700 calls are to be served. This is far in excess of what would be considered permissible in practice, but it will serve for the illustration. The occupancy of the translator will be $y=0.35\times 7{,}700/3{,}600=0.75$ erlang.

From equation 5(A) the average delay $m/T=1.5$ so $m=0.525$ second. From equation 3(A) $P(>0)=0.75$, that is 75%, or 5,775, of the calls will be delayed. The average delay of a delayed call from equation 6 will be $M=0.7$ second.

The gate occupancy from equation 8 is $G=0.5625$ erlang, or 2,025 seconds. The average number of calls waiting during delays, computed from equation 9 is $h=2$. According to equation 10, the average number of calls entering the gate will be the same, $H=2$.

Since the total number of delayed calls D is 5,775, the number of gate operations in 1 hour k, from equation 12 is 2,887.5, and the mean duration of a gate holding time g from equation 13 is 0.7 second. The average delay outside, as well as the average delay inside the gate will be $g/2$ or 0.35 second.

Next the probability may be asked that a gate will be held for at least say, 2.1 seconds, or six translator holding times. The charts show that if there is one channel with an occupancy of 0.75, there is a probability of 0.04, and a call will be delayed at least six holding times, but since only 75% of the calls are delayed, the probability of a gate holding time of at least 2.1 seconds is $Q(>t)=0.0535$. In other words, roughly 5% of the gate holding times, or 155, will be over 2.1 seconds.

If the question is asked as to the probability that a call will have to wait at least 1 second to enter the gate, the answer is likewise a little over 5%, say 6%, according to the suggestion made at the end of the section "Delays Outside the Gate," since the probability of a waiting time for the gate of over 1.05 seconds is equivalent to the probability of gate holding times over 2.1 seconds.

The probability that more than 6 calls will be admitted to the gate is the same, $R(>6)=0.0535$, so it will happen 155 times in 1 hour. This is derived from equation 16, where $R(>n)=Q(>t)$, if $n=Nt/T$.

If the translator is accessed from s sources, say, directors, which are arranged in a preferential chain, the highest ranking source may expect an average delay (if it is delayed at all) of 0.35 second, and the lowest ranking source 1.05 second; this is derived from equation 17.

The probability of gate holding times which will be exceeded, on the average, once in a busy hour may now be asked. There is, of course, a probability of still longer gate holding times, but as they may occur only once in several busy hours, these can be disregarded in the present considerations. Since there are 2,887.5 gate closures in each busy hour, and $P(>0)=0.75$, the charts should be consulted for the delays which are exceeded with strict sequence service with a probability of $0.75/2{,}887.5=0.00026$. Extrapolation of the charts gives this probability at $16T$, or 5.6 seconds. This means that on the average, once in each busy hour, the gate holding time will be over 5.6 seconds, and that more than 16 calls will be admitted into the gate.

If there are 17 calls in the gate, the lowest ranking source must wait 5.95 seconds for service, and $5.95+5.6+0.35=11.9$ seconds is about the longest delay in a busy hour, if it is assumed that 17 calls are admitted to the gate following a gate occupancy with 16 calls, that the probability of more than 17 calls can be disregarded, that the source turning in the earliest request happened to be the lowest ranking of the 17, and that the holding time for servicing the call is added.

As the last question on this example, the probability is asked that a call will not be delayed for over 3 seconds. From previous data it is known that this probability will be on the order of 0.01; therefore a multiplier $K=1.2$ according to the section "Distribution of Delays" is chosen and in the charts is found for "$N=1$, $y/N=0.75, t/T=3\times 1.2/0.35$," by extrapolation $P(>t)=0.005$, or that 38.5 calls in the busy hour will be delayed for at least 3 seconds or that 7,661.6, or 99.5%, calls will be served within 3 seconds.

EXAMPLE 2

As a second example, a team of 10 toll operators may be considered, receiving a total of 170 calls in 1 hour, each requiring 120 seconds of work by an operator. This example is more representative than the first of actual conditions, and comparison of the two will show how the results differ when the same gate principle is used in so widely divergent applications.

Now, each operator considered as a channel, the total traffic carried by the 10 channels will be $170\times 120/3{,}600=5.67$ erlangs. The occupancy of each operator will be 0.567 erlang. From the charts an average delay of $mT=0.011$ is obtained, so $m=1.32$ seconds. The charts also give for the probability that a call will be delayed $P(>0)=0.06$. This means that only 6% of the calls, or 10.2 altogether, will be delayed at all. The average delay of these calls will be $M=m/P(>0)=22$ seconds.

The gate occupancy from equation 8 is $G=0.034$ erlang, or 122.4 seconds. The average number of calls waiting during delays, computed from equation 9 is $h=1.83$, which is also H, the average number of calls admitted to the gate.

The total number of delayed calls D, is 10.2, therefore the expected number of gate operations k, in 1 hour will be 5.57, from equation 12, and the mean duration g of a gate holding time, from equation 13, 22 seconds. The average delay outside the gate as well as inside will be $g/2$, or 11 seconds.

Next, it may be inquired if there is the probability that the gate will be held for at least 24 seconds, or one-fifth of the operator's work time. The charts show that there is a probability of 0.02 that a call will be delayed for at least one-fifth of its holding time if there are 10 channels with an occupancy of 0.567; but since only 6% of the calls will be delayed at all, the probability of a gate holding time of at least 22 seconds will be $Q(>t)=0.02/0.06=0.333$. In other words, one-third of the gate holding times, or 1.86, will be over 24 seconds.

The probability that a call will wait at least 12 seconds to enter the gate will likewise be one-third. The probability that at least $24\times 10/120=2$ calls will be admitted to the gate is the same, and it will happen 1.86 times in 1 hour. If the probability is asked for that at least four calls will be admitted to the gate, the answer is 0.05, or once in $1/(0.05\times 5.57)=3.6$ hours.

If the operators are accessed from s trunks, arranged in a preferential chain, the highest ranking trunk may expect an average delay (if it is delayed at all) of 11 seconds; the lowest ranking trunk, 33 seconds. However, this spread is relatively unimportant. First, only 6% of the calls enter a gate and, second, two-thirds of the gate occupancies will hold only one or two calls. When a call is not delayed, or when it is alone in the gate, the rank of the trunk is irrelevant. When two calls are in the gate (which happens less than twice in an hour), the average spread between the two will be $T/N=12$ seconds, and only when three calls are in the gate will the average spread between the ranks exceed this. As has been seen, only very rarely will there be four or more calls in the gate.

As the last question on this example, the proportion is asked of all calls waiting longer than 10 seconds, or one-twelfth of the operators' work time. From the charts for 10 channels and an average occupancy of 0.567, with strict sequence service the probability of delays at least one-twelfth of the holding time is 0.045, that is, 4.5% of the calls will be delayed over 10 seconds. According to the section "Distribution of Delays," in such cases a multiplier $k=1$ is satisfactory, and there should be no difference between gate sequence and strict sequence service. As a matter of fact, in the exponential holding-time case with 10 channels, 0.567 occupancy and 120 seconds' average holding time, fewer calls are delayed over 10 seconds if the calls are answered at random than if they are answered in the sequence of their origination. The proportion of 4.5% calls delayed over 10 seconds conforms with the general operating rule that 95% of the calls should be answered within 10 seconds.

The curious aspect of this problem is that while only 6% of the calls are delayed at all, three-fourths of these, that is, 4.5% of all calls, are delayed for over 10 seconds. As the average delay of delayed calls is 22 seconds, it is evident that the delays, as far as they exist, are substantial. But the number of calls under consideration is very small: only about 10 calls have any amount of delay (although about eight of these have delays of over 10 seconds).

Comparing the two examples results in several conclusions which are immediately apparent. First, in the second example all the delay times are much longer than in the first one, because of the great difference in the holding times: 120 versus 0.35 second. Nevertheless, the difference in the delays is not as great as the relative magnitudes of the holding times would suggest. Second, in the second example only a very minor portion of the calls are delayed and what happens to them has relatively little bearing on the over-all problem, while in the first example three calls out of four are delayed and their efficient handling determines the efficiency of the system. The difference here is because in the first case there is only one channel, while in the second case there are 10, and at the same time the average occupancy in the first case is relatively high and in the second case relatively low. Third, for the same reason the number of calls handled by the gate in the second case is very low (it should hardly ever exceed four, and in most cases will be only one or two) while in the first example there is a chance that once in an hour 16 or more calls will be in the gate, even though the average number of calls in the gate is about the same as in the second example.

From the preceding facts one might ask the purpose of providing in the second case any gate arrangement at all, while in the first case it is obviously of the greatest importance to the operation of the system. A few seconds of delay occasionally is acceptable in translator operations; however, the fact that each hour a fair number of calls is expected to exceed, say, 10 times the average holding time indicates that the margin against unexpected overloads is small. Therefore the distribution of delays provides a measure of the permissible occupancy of the channel. In the second example only about eight calls in each hour are expected to be delayed for over 10 seconds, and even delays of 30 seconds or more, (which is many times longer than the longest practically occurring delays in the translator system) are relatively short compared with the work time of the operator, so it appears that here is plenty of margin for unexpected overloads. However, it must be remembered that in the second example the delays are those experienced by a customer while waiting for an operator, and some customers may consider over 30 seconds as excessive, even if it occurs only twice in an hour, as in our example, and in spite of the fact that it will take on the average 120 seconds for the operator to process the call. Therefore an efficient arrangement, which makes it possible to cut down excessive delays, also has full justification in the second application.

References

1. PROBABILITY AND ITS ENGINEERING USES (book), T. C. Fry. D. Van Nostrand and Company, Inc., New York, N. Y., 1928, pp. 373–87.

2. APPLICATION OF THE THEORY OF PROBABILITY TO TELEPHONE TRUNKING PROBLEMS, E. C. Molina. *Bell System Technical Journal*, New York, N. Y., vol. 6, July 1927, p. 461.

3. *Trunking Report No. 54*, British Post Office, London, England, 1936.

4. MOE'S PRINCIPLE (book), Arne Jensen. The Copenhagen Telephone Company, Copenhagen, Denmark, 1950.

5. WORKING CURVES FOR DELAYED EXPONENTIAL CALLS SERVED IN RANDOM ORDER, Roger I. Wilkinson. *Bell System Technical Journal*, New York, N. Y., 1953, pp. 360–83.

6. DELAY PROBABILITY CHARTS FOR TELEPHONE TRAFFIC WHERE THE HOLDING TIMES ARE CONSTANT, Imre Molnar. Automatic Electric Company, Chicago, Ill., 1952.

A Precision Torque Balance for the Measurement of Small Torques

R. J. MATTHYS
NONMEMBER AIEE

L 74
1
55

THE development and testing of torque motors for precision gyroscopes require the accurate measurement of the small torques developed. These are stall torques, and are measured at various angular positions of the rotor with respect to the stator. The magnitude of these torques is generally less than 50,000 dyne-centimeters (dyne-cm).

The primary problem in a measuring device for this purpose is the friction level in the pivots or bearings of the device. In the past, jewelled-pivot, ball-bearing, and fluid-bearing types have been used; but each has its disadvantages. The ball-bearing type is rugged but has a high (20-dyne-cm) friction level. Jewelled-pivot types have a 15-dyne-cm uncertainty level but are subject to considerable breakage because of the delicate pivot construction. Fluid-bearing types are quite rugged and can be made with an uncertainty level of 1 dyne-cm, but they entail elaborate fluid supply and filtering problems. A knife-edge pivot cannot be used, as the pivot must be stable to eliminate response to radial forces.

A torque-measuring device with a threshold of 0.05 dyne-cm using a crossed-reed flexure pivot is described in this paper. This type of pivot is simple, fairly rugged, and requires no maintenance.

Principles of Operation

In principle, the torque balance acts as an angular spring. When a torque is applied to the balance, the beam rotates from the zero torque or null position to a position where the spring torque generated by rotation of the beam exactly counterbalances the applied torque. To minimize errors, the maximum angle of the beam rotation is kept small, less than 0.004 degree. The torque balance is shown schematically in Fig. 1. The amplifier and permanent magnet solenoid produce a torque proportional to the displacement angle and thus act as an electrical spring. The applied torque is measured electrically by measuring the current into the solenoid.

The total beam spring constant consists of the electrical spring stiffness of the amplifier circuit plus the stiffness of the crossed-reed pivot, the elastic restraint of the capacitor pickoff, top or bottom heaviness of the beam itself, and the change in the floated volume of the fluid dampers as the beam rotates. All of these stiffnesses are small with respect to the electrical stiffness except that of the pivot, which uncompensated is 1.5×10^6 dyne-cm per radian.

The balance is shown in block diagram form in Fig. 2 for analysis. Using well-known principles of servomechanism theory, the transfer function of electric torque output to torque input is as follows

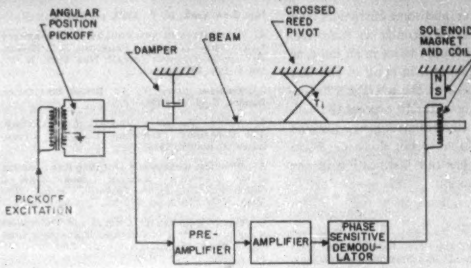
Nomenclature

t_i = instantaneous input torque
T_i = constant input torque
θ = angle of beam rotation
t_o = instantaneous electric output torque
T_o = constant electric output torque
T_o' = constant output torque with K or C nonlinear
J = beam polar moment of inertia
$C = C_1 + C_2 + C_3 + C_4$ = sum of all beam stiffnesses except the electrical stiffness
K = electrical spring stiffness of the beam
F = beam damping
C_1 = crossed-reed pivot stiffness
C_2 = pickoff negative elastic restraint
C_3 = beam stiffness from being top heavy or bottom heavy
C_4 = beam stiffness caused by change in the floated volume of the fluid dampers as the beam rotates
τ = time constant of demodulator filter circuit
E = Young's modulus of elasticity
$I = (bt)^3/12$ = cross-sectional moment of inertia of the reeds on each side of the pivot
V_t = vertical tension force on the crossed-reed pivot
V_c = vertical compression force on the crossed-reed pivot
H = horizontal force on the crossed-reed pivot
T_{read} = torque reading on potentiometer
L = length of reeds
b = width of reeds
t = reed thickness
s = displacement of pivot center of rotation with rotation
p = operator $d/(dt)$

Paper 55-486, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 30, 1955; made available for printing June 9, 1955.

R. J. MATTHYS is with the Minneapolis-Honeywell Regulator Company, Minneapolis, Minn.

Credit is given to K. H. White for his work on the mechanical design of the balance. Credit is also given to E. P. Blair for fabricating and assembling the balance.



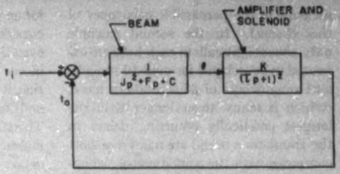


Fig. 1 (left). Schematic diagram, torque balance

Fig. 2 (above). Block diagram, torque balance

$$\frac{t_o}{t_i}=\frac{K}{(Jp^2+Fp+C)(\tau p+1)^2+K} \tag{1}$$

The amplifier time constant is small and can be ignored, and since only the steady state value is wanted

$$\frac{T_o}{T_i}=\frac{K}{C+K} \tag{2}$$

Equation 2 shows that if the electrical stiffness is much greater than the rest of the beam stiffnesses, the accuracy of the balance will be almost independent of small changes in any of the beam stiffnesses.

Wittrick[1] has shown that the crossed-reed pivot spring constant can be tripled, made zero, or even negative, by applying horizontal or vertical forces to the pivot through the center of rotation. Wittrick's curves of pivot stiffness for a 90-degree reed intersection angle as a function of the horizontal and vertical forces applied are given in nondimensionalized form in Fig. 3. At point *A*, pivot stiffness is zero with only a vertical tension force applied.

At *A* the tension force is

$$V_t=\frac{8EI}{(L)^2} \tag{3}$$

Four beryllium copper reeds, 3 inches long, 0.250 inch wide, and 0.015 inch thick, and intersecting at 90 degrees were used. The required vertical tension force with these dimensions is 930 grams. A cylindrical weight, slotted to clear the reeds, was added to the beam at the center of rotation to make the total beam weight 930 grams. This weight was trimmed slightly to compensate for the negative elastic restraint of the capacitor pickoff and the positive stiffness of the fluid dampers.

Measurement Technique

The balance has three ranges of operation: 0.05 to 2 dyne-cm; 2.00 to 1000 dyne-cm; and 1,000 to 100,000 dyne-cm. The first two ranges are read directly; the third requires a multiplier of 10.

A potentiometer circuit is used to measure the current in the solenoids, and torques up to 1,000 dyne-cm are read directly from the potentiometer. Torques over 1,000 dyne-cm are measured by balancing off all of the torque exceeding 1,000 dyne-cm with accurate weights in the weight pan and reading the remainder on the potentiometer.

To avoid the difficulty of balancing the beam to 0.00±0.05 dyne-cm with no input torque, the beam is balanced approximately to 0 dyne-cm. This near-zero reading is subtracted from the reading which is obtained when torque is applied.

Mechanical Features

The beam structure is shown in Fig. 4, and the balance in Figs. 5 and 6. The main mechanical features of the balance are the crossed-reed pivot, a capacitor-type position pickoff, two symmetrically opposed permanent magnet solenoids, a weight pan suspended from a double knife edge,[2] two symmetrically opposed fluid dampers, and two rotating chucks centered on the axis of rotation.

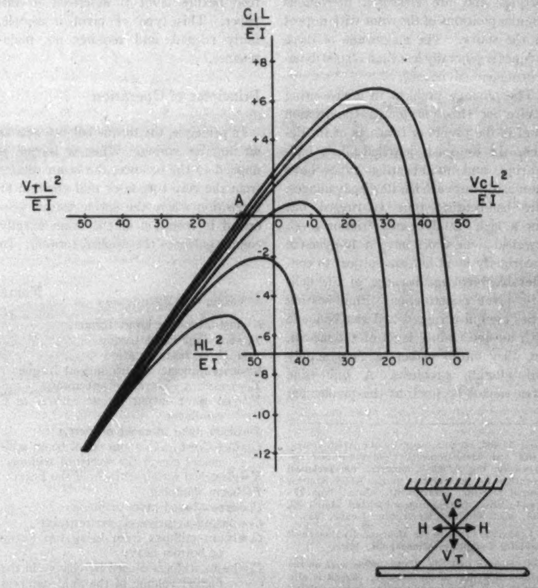


Fig. 3. Crossed spring pivot stiffness as a function of horizontal and vertical forces

The reeds were specially heat-treated to insure flatness and to avoid oilcan effects. They are insulated at the ends with micalex strips so that they can be used as flex leads to the solenoid coils and the middle plate of the capacitor pickoff on the beam.

The capacitor-type pickoff has essentially infinite resolution, a negative elastic restraint of 2.8×10^4 dyne-cm per radian, which is due entirely to loading effects, and a high sensitivity of 6,000 volts per radian. Considerable effort was expended on its parts to attain a high degree of mechanical stability, as the pickoff displacement is only 1.3×10^{-8} inch for a 0.05-dyne-cm torque. A disadvantage of the pickoff is its high output impedance which is an inverse function of frequency.

The two permanent magnet solenoids are each linear to 0.4 per cent but, when used in the manner shown in Fig. 4, their combined linearity increases to 0.05 per cent because their nonlinearities tend to cancel each other. The weight pan is hung from a double parallel knife-edge suspension so that it always hangs vertically from the top knife edge even though weights are not centered on the pan. The weight pan is used to calibrate the balance and to measure torques over 1,000 dyne-cm.

Two fluid dashpot dampers were added to the beam, even though the balance is stable without them, to cut down ambient vibration effects. The weight pan dashpot also damps out the swinging of the weight pan on its knife-edge suspension. Symmetry of the two dashpots allows the buoyancy (4,200 dyne-cm) and the surface tension (33.5 dyne-cm) torques of each damper to cancel out although the changes in buoyancy with rotation do not. The changes in buoyancy with rotation produce a positive stiffness of 1,600 dyne-cm per radian.

The two rotating chucks, one on the front of the balance and one on the back, are used to mount the torque motors on the balance. Each chuck will rotate through ±15 degrees and has radial and rotational centering adjustments. For production work a standard torque motor is mounted in the chuck on the rear side so that the difference between the torque of the motor being tested and the desired or standard value can be measured directly.

Electrical Features

The main electric components of the torque balance are the preamplifier, amplifier, ring-type phase-sensitive de-

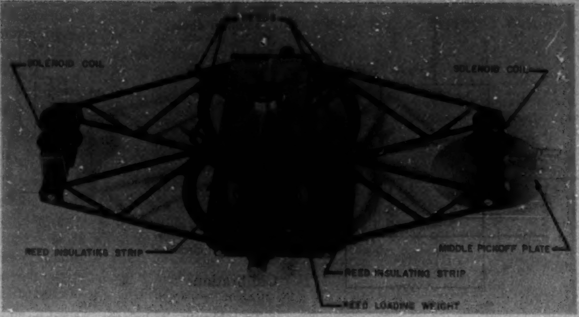

SOLENOID COIL
SOLENOID COIL
REED INSULATING STRIP
MIDDLE PICKOFF PLATE
REED INSULATING STRIP
REED LOADING WEIGHT

Fig. 4. Beam structure

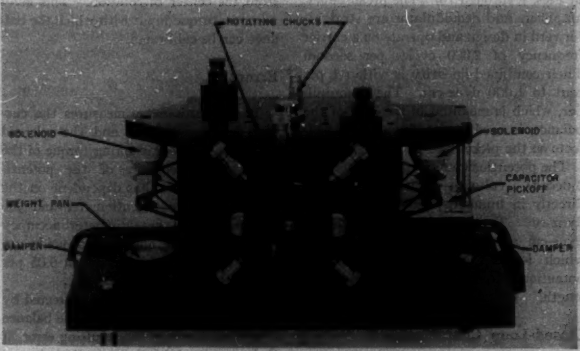

ROTATING CHUCKS
SOLENOID
SOLENOID
CAPACITOR PICKOFF
WEIGHT PAN
DAMPER
DAMPER

Fig. 5. Torque balance

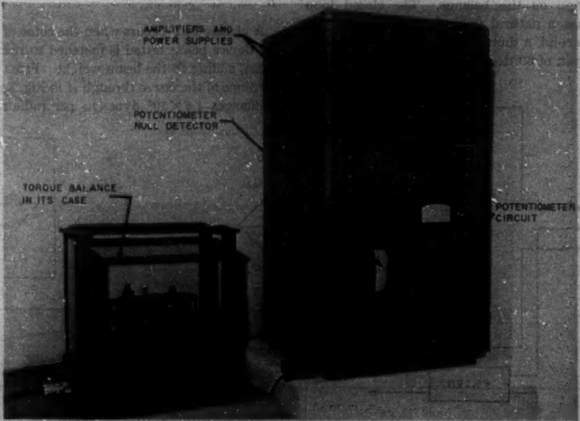

AMPLIFIERS AND POWER SUPPLIES
POTENTIOMETER NULL DETECTOR
TORQUE BALANCE IN ITS CASE
POTENTIOMETER CIRCUIT

Fig. 6. The complete balance system

l, 74
1
55

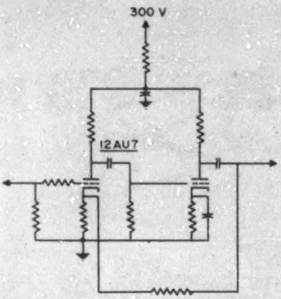


Fig. 7 (left). Preamplifier circuit

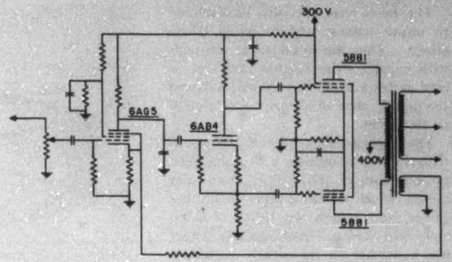


Fig. 8 (right). Amplifier circuit

modulator, and potentiometer circuit. Brief schematics of these units are shown in Figs. 7 through 10. The preamplifier, amplifier, and demodulator are straightforward in design and operate on a carrier frequency of 2,600 cycles per second. Their combined linearity is within 1 per cent to 1,000 dyne-cm. The preamplifier, which is mounted in the base of the balance, is used to cut down loading effects on the pickoff.

The potentiometer circuit consists of a 3-decade voltage divider calibrated directly in hundreds, tens, and units of dyne-cm. Tenths and hundredths of dyne-cm are read on the null detector which is an automatic self-balancing potentiometer having a 30-inch scale length.

Closed-Loop Characteristics

As Fig. 2 shows, the amplifier and permanent magnet solenoids form a closed loop on the balance. The closed loop has a natural frequency of 6 cycles per second, a damping ratio of 2, and a loop gain of 60 decibels.

Calibration

To calibrate the balance, an accurately known weight is placed on the weight pan which is hung at a known distance from the center of rotation. This gives a known torque input with which the balance can be calibrated.

Errors

The potentiometer measures the current in the solenoids, and the solenoid torque is the electric spring torque of the beam. The accuracy of the potentiometer reading is thus dependent on the linearity of the potentiometer and the solenoids. The potentiometer accuracy is 0.05 per cent and the solenoids are linear to 1,000 dyne-cm within 0.05 per cent.

Changes in K or C can be corrected by recalibrating the balance. If the balance is not recalibrated, the resulting error, if C changes by ΔC (see Appendix I) is

$$\frac{\Delta C}{K} T_t \tag{4}$$

A change in C occurs when the rotor of the device being tested is fastened to the beam, adding to the beam weight. From the slope of the curve through A in Fig. 3, C changes 1.8×10^3 dyne-cm per radian per gram change in beam weight. With K being 1.5×10^7 dyne-cm per radian, an 8.5-gram change in beam weight will produce an error of 0.1 per cent.

If the balance is not recalibrated when K changes by ΔK, the resulting measurement error is

$$\frac{C\Delta K}{K(K+\Delta K)} T_t \tag{5}$$

With a value of C equivalent to an 8.5-gram change in beam weight, a 20-per-cent change in K produces an error of 0.05 per cent for $K=1.5\times10^7$ dyne-cm per radian.

Errors are introduced if either K or C vary with beam rotation. If K changes at some angle θ_1 from the slope at null by $\Delta_1 K$ and the input torque is large enough to rotate the beam to some angle θ greater than θ_1, the torque error can be approximated by

$$\frac{C\Delta_1 K}{K}(\theta-\theta_1) \tag{6}$$

If the worst case is picked, i.e., θ_1 is zero, the torque error becomes the same as for the linear case

$$\frac{C\Delta_1 K}{K^2} T_t \tag{7}$$

By the same procedure, the maximum torque error caused by C_1 being nonlinear

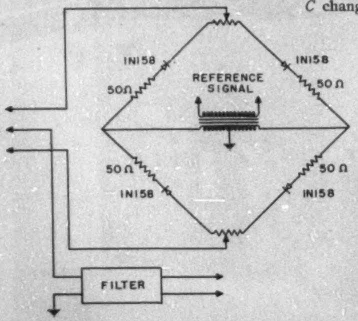


Fig. 9. Simplified phase-sensitive demodulator circuit

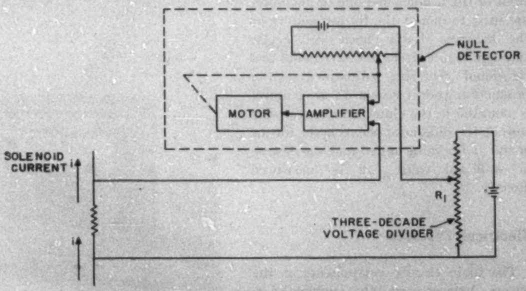


Fig. 10. Potentiometer circuit for measuring solenoid current

is also the same as for the linear case

$$\frac{\Delta C_1}{K} T_i \tag{8}$$

The maximum change in C with rotation was measured as 1.5×10^4 dyne-cm per radian, which gives a maximum error of 0.1 per cent.

The beam rotates off null when torque is applied, and this introduces an error in the angular position at which torque is measured. The beam has a high stiffness to keep this error small. The maximum error resulting is 0.004 degree and occurs for torques of 1,000 dyne-cm or larger.

A characteristic of the crossed-reed pivot is that its center of rotation moves with rotation.[3] Young's equation for this movement is

$$s = (0.4133\theta - 0.012258\theta^2 + 0.001267\theta^3 - 0.0000075\theta^4)10^{-3}L$$

For a maximum angle of 6.7×10^{-3} radian, the center of rotation moves the negligible amount of 8.3×10^{-6} inch.

There is one other effect that is more of a nuisance than an error. This is the drift of the pickoff null with temperature. This drift is proportional to temperature and to the nonelectrical beam stiffnesses and is independent of the servo loop gain. By making temperature variations and the nonelectrical portion of the total beam stiffness small and by temperature-compensating the pickoff, the drift has been reduced to 0.3 dyne-cm per hour or less. Error caused by this drift in torque reading cancels out when the zero-torque reading is subtracted from the reading with torque applied, as both drift the same amount.

Accuracy

Over-all accuracy of the balance depends primarily on the accuracy of calibration and on solenoid and potentiometer linearity and, secondly, on the constancy of reed stiffness with rotation. All other errors are reduced to negligible proportions by the large electric spring constant. The accuracy of torque measurements made on the balance is 0.3 per cent+0.05 dyne-cm for the range of 0.05 to 2 dyne-cm, 0.3 per cent+0.1 dyne-cm for the range of 2 to 1,000 dyne-cm, and 0.1 per cent+15 dyne-cm for the range of 1,000 to 100,000 dyne-cm.

Conclusion

By the use of principles and devices described herein it has been possible to make an accurate and fairly rugged torque-measuring device with a very low uncertainty level for the measurement of small torques in the range of 0.05 to 100,000 dyne-cm.

Appendix I. Torque Error Caused by a Change ΔC in C

$$T_o = \frac{K}{C+K} T_i$$

Since the potentiometer circuit is calibrated directly in dyne-cm of input torque

$$T_{\text{read}} = T_i = \frac{C+K}{K} T_o$$

If C changes by ΔC

$$T_o' = \frac{K}{C+\Delta C+K} T_i$$

$$T_{\text{read}} = \frac{C+K}{K} \quad T_o' = \left(1 - \frac{\Delta C}{C+\Delta C+K}\right) T_i$$

If $C+\Delta C << K$

$$T_{\text{read}} = T_i - \frac{C}{K} T_i$$

The measurement error in dyne-cm is then

$$\frac{\Delta C}{K} T_i$$

Appendix II. Torque Error Caused by a Change ΔK in K

$$T_o = \frac{K}{C+K} T_i$$

$$T_{\text{read}} = T_i = \frac{C+K}{K} T_o$$

If K changes by ΔK

$$T_o' = \frac{K+\Delta K}{C+K+\Delta K} T_i$$

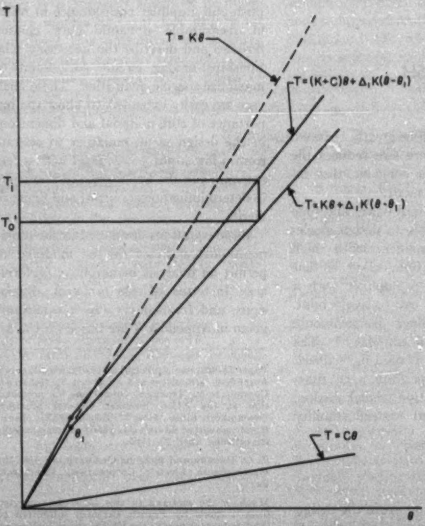


Fig. 11. Calculation of torque error caused by K being nonlinear

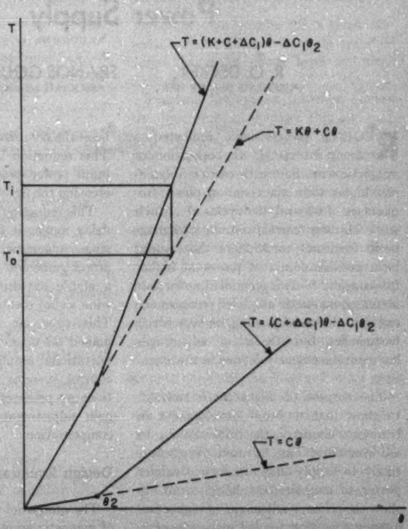


Fig. 12. Calculation of torque error caused by C being nonlinear

$$T_{\text{read}}=\frac{C+K}{K}\ T_o'=\frac{C+K}{K}\left(\frac{K+\Delta K}{C+K+\Delta K}\right)T_t$$
$$=\left(1+\frac{C\Delta K}{K(C+K+\Delta K)}\right)T_t$$

If $C<<K+\Delta K$

$$T_{\text{read}}=T_t+\frac{C\Delta K}{K(K+\Delta K)}\ T_t$$

The measurement error in dyne-cm is

$$\frac{C\Delta K}{K(K+\Delta K)}\ T_t$$

Appendix III. Torque Error Caused by K Being Nonlinear

To simplify the mathematics, the nonlinearity of K is approximated by two straight lines, as shown in Fig. 11

$$T_o'=\frac{K\theta+\Delta_1K(\theta-\theta_1)}{(K+C)\theta+\Delta_1K(\theta-\theta_1)}T_t$$

$$T_{\text{read}}=\frac{C+K}{K}\ T_o'=\frac{C+K}{K}\times$$
$$\left[\frac{K}{C+K}+\frac{C\Delta_1K(\theta-\theta_1)}{(C+K)[(C+K)\theta+\Delta_1K(\theta-\theta_1)]}\right]T_t$$

$$T_t=(C+K)\theta+\Delta_1K(\theta-\theta_1)$$

$$T_{\text{read}}=T_t+\frac{C\Delta_1K}{K}(\theta-\theta_1)$$

The torque error is then

$$\frac{C\Delta_1K}{K}(\theta-\theta_1)$$

Appendix IV. Torque Error Caused by C_1 Being Nonlinear

To simplify the mathematics, the nonlinearity of C_1 is approximated by two straight lines, as shown in Fig. 12.

$$T_o'=\frac{K\theta}{(K+C+\Delta C_1)\theta-\theta_2\Delta_1C}\ T_t=\left[\frac{K}{C+K}-\right.$$
$$\left.\frac{K\Delta C_1(\theta-\theta_2)}{(C+K)[(K+C+\Delta C_1)\theta-\theta_2\Delta C_1]}\right]T_t$$

$$T_{\text{read}}=\frac{C+K}{K}\ T_o'=T_t-$$
$$\frac{\Delta C_1(\theta-\theta_2)}{(K+C+\Delta C_1)\theta-\theta_2\Delta C_1}\ T_t$$

$$T_t=(K+C+\Delta C_1)\theta-\theta_2\Delta C_1$$

$$T_{\text{read}}=T_t-\Delta C_1(\theta-\theta_2)$$

The torque error is then

$$\Delta C_1(\theta-\theta_2)$$

References

1. **THEORY OF SYMMETRICAL CROSSED FLEXURE PIVOTS, W. H. Wittrick.** ***Australian Journal of Scientific Research*****, East Melbourne, Australia, vol. I, no. 2, 1948, pp. 121–34.**

2. **A KNIFE EDGE BALANCE FOR WEIGHINGS OF THE HIGHEST ACCURACY, F. A. Gould.** ***Proceedings*****, Physical Society, London, England, vol. 62, pt. 12-B, Dec. 1, 1949.**

3. **AN INVESTIGATION OF THE CROSS SPRING PIVOT, W. E. Young.** ***Journal of Applied Mechanics*****, New York, N. Y., vol. 11, no. 2, June 1944, pp. 113–20.**

4. **ELECTRICAL MEASUREMENTS (book), F. K. Harris. John Wiley and Sons, Inc., New York, N. Y., 1952, pp. 150–85.**

5. **THE PROPERTIES OF CROSSED FLEXURE PIVOTS, W. H. Wittrick.** ***Aeronautical Quarterly*****, London, England, vol. 2, Feb. 1951, pp. 272–92.**

6. **DESIGN CHARACTERISTICS OF CROSS SPRING PIVOTS, L. W. Nichols, W. L. Wunsch.** ***Machinery*****, London, England, vol. 79, Oct. 11, 1951, pp. 645–51.**

Operational Magnetic Amplifier with Audio-Frequency Transistor Power Supply

R. O. DECKER
ASSOCIATE MEMBER AIEE

FRANCIS GOURASH
ASSOCIATE MEMBER AIEE

RECENT papers have indicated a strong interest in the operation of magnetic amplifiers with carrier frequencies higher than the common power frequencies of 60 and 400 cycles.[1,2] Much work has been carried out, e.g., on magnetic frequency multipliers that would be a possible source of power at higher frequencies.[3,4] It is generally known that increased magnetic amplifier performance can be achieved by using higher carrier frequencies, but the exact advantages have not been clearly defined in the literature.

The purpose of this paper is twofold: to show that a recent development in transistor circuitry, the transistor d-c to a-c converter, can be used very effectively to supply audio-frequency carrier power to magnetic amplifiers; and, to show the primary advantage of higher frequency operation for full-wave magnetic amplifiers is that for a given power output level the core size can be greatly reduced. This reduction in core size reduces the input power and time constant, thus increasing the figure of merit.

This reduction in input power and time delay makes it feasible to design single-stage magnetic amplifiers with high power gains and low time delays, so that a stable low-drift d-c amplifier with a wide signal passband can be easily built. This type of amplifier is commonly named an operational amplifier.[5,6] The operational amplifier is used in feedback control systems to perform such functions as preamplification, signal mixing, gain adjustment, and system stability compensation.

Design Equations

The quantities of interest in the design of magnetic amplifiers are time constant, control power, power gain, and figure of merit. The figure of merit is defined as the ratio of power gain to time constant. Approximate equations for these quantities have been developed previously.[7,8] These analyses which include the effects of circuit parameters describe push-pull amplifier performance in terms of measurable dynamic core characteristics and describe the necessary relationships among circuit parameters to maximize specific quantities. These analyses are easily extended to show the importance of core material and dimensions in the design of an amplifier to operate from a low input power level with a fast response and the importance of frequency in determining power output and figure of merit.

Basic equations developed in the aforementioned analyses can be modified to permit an analysis of amplifier performance in terms of core material, dimensions, and frrquency. These equations, given in Appendix I, are based on the as-

---

Paper 55-463, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 24, 1955; made available for printing April 20, 1955.

R. O. DECKER and FRANCIS GOURASH are with the Westinghouse Electric Corporation, Pittsburgh, Pa.

Much of the material in this paper is excerpted from a thesis by Francis Gourash submitted to the Graduate School of the University of Pittsburgh in January 1955 in partial fulfillment for the requirements for the degree of Master of Science.

sumption that the change in preset core flux level is proportional to the change in net average magnetomotive force (mmf) during the flux-setting half-cycle, e.g., $\Delta\phi = KN_c\Delta I_c$ where $\Delta\phi$ is the change in the preset flux level caused by the change in control ampere-turns $N_c\Delta I_c$. This expression describes the locus of reset points of the minor loops actually followed by the core during each cycle.[9] K, which is the slope of this locus, can be expressed in terms of an effective ΔB and ΔH. These quantities, as defined, are shown in Fig. 1. In push-pull magnetic amplifier circuitry the biasing point for the cores and the interaction through load circuitry limit the net amount of mmf required to control a core to approximately one half of the range shown in Fig. 1.

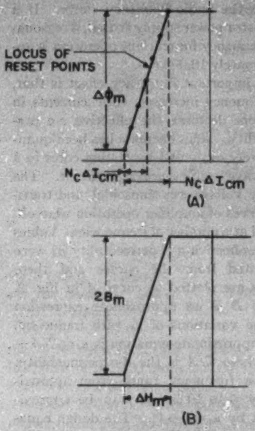


Fig. 1 (left). Control curve showing ΔH_m

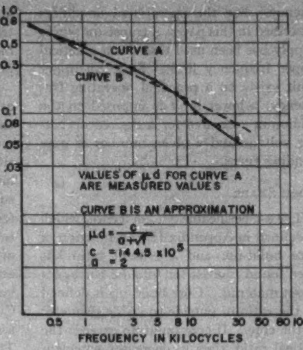


Fig. 2 (above). Experimental curve of a-c permeability versus frequency for 1-mil Hymu 80 core material

Design Parameters

The modified equations given in Appendix II have been developed to show the effects of core materials, core dimensions, and frequency on amplifier performance. The equations are helpful in pointing out the effects of these parameters on control power, time constant, and figure of merit.

CORE MATERIALS

The amount of maximum control power is a specified quantity in many control applications. If low drift is to be realized, this maximum control power should be that given by equation 7, i.e., the maximum control power available should be the same as the maximum control power needed to control the core. In this way, the amplifier drift in terms of input power is a smaller percentage of the maximum signal power.

Thus, ΔH and β are the variables when designing for a given input power level. β, which relates N_c^2/R_c to the winding area, should be as large as possible, i.e., the core should be wound with a large amount of control copper. However, a limit is imposed on β because of winding difficulties. ΔH, which is a measure of core sensitivity, should be small, i.e., a sensitive core material such as Hymu *80* or Supermalloy is desirable for low input power.

CORE DIMENSIONS

If maximum control power and response time are specified for a given application, the product of input power and response time restrict the volume of the core. This product, given as equation 9 is

$$(\Delta P_{in})(T) = C_2(\Delta B)(\Delta H)V \text{ watt-seconds}$$

Thus, the core volume for a given material may be selected from the specified values for input power level and time constant.

Practical limitations, however, govern the dimensions of the core that make up the required volume. The area of the hole required to permit the winding shuttle to pass for machine-wound cores places a limitation on the minimum value of l_m. In addition, l_m must be large enough to allow sufficient gate winding turns to be wound on the core to produce output power and reduce the exciting current to a small value while maintaining the required β for the input power level. A physical limitation on minimum area of the core is that a minimum of four turns of core material are required in the core to eliminate the effects of an air gap, and the thinnest toroidal materials available are made of 1/8-mil strip. The minimum height of the core is restricted to 1/8-inch since this is the smallest strip width available, while the radius of the winding shuttle restricts the maximum height of the core.

Another limitation on the core area exists because the operating voltage is proportional to the core area and the fre-

---

Nomenclature

ϕ_s = saturation flux level, maxwells
$\Delta\phi$ = change in preset flux level of core
B_m = saturation flux density, gausses
ΔB = change in flux density
ΔH = change in magnetizing force, oersteds
μ_d = effective a-c permeability of core
l_m = mean magnetic path length, centimeters (cm)
A_c = area of iron in the core, cm^2
V = volume of core, cm^3
O.D. = outside diameter of the core
I.D. = inside diameter of the core
r = the ratio of outside diameter to inside diameter
I_c = control current
$\Delta I_{cm}'$ = value of maximum control current for push-pull circuit operation
ΔI_{cm} = value of maximum control current for single core operation
ΔI_L = change in output load current
K_I = current gain over linear region = $\Delta I_L/\Delta I_c$
ΔP_{in} = input control power
ΔP_{out} = output power
K_p = power gain over the linear region, $\Delta P_{out}/\Delta P_{in}$
T = time constant, seconds (linear region)
M = figure of merit per second over linear region, K_p/T
N_c = number of turns in control winding
N_g = number of turns in gate winding
N_b = number of turns in bias winding
N_{FB} = number of turns in feedback winding
α = portion of window area taken by the gate winding
β = portion of window area taken by control winding
E_m = maximum value of supply voltage
$V_{M(n)}$ = voltage across mixing resistors 1 or 2
$V_{G(n)}$ = voltage across gate winding of cores A, B, C, or D
V_L = output voltage across load resistor
f = frequency, cycles per second (cps)
R_c = control circuit resistance
R_m = mixing resistor
R_L = output load resistor
R_F = rectifier forward resistance
R_w = gate winding resistance
R_g = internal resistance of amplifier ($R_w + R_F$)
$R' = 2R_g + R_L + (R_L R_g)/R_m$
R_B = bias resistor
R_{FB} = feedback resistor
$K_{0(n)}$, $K_{0(n')}$ = constants of proportionality

quency. For the operational amplifiers described in this paper, a transistor power supply has been used that has a practical upper frequency limit. Since an upper limit exists on a practical operating frequency, a lower limit is imposed on the minimum core area for a desired operating voltage and a given number of gate winding turns.

CORE SHAPE

The a-c permeability μ_d of a core, which is defined in Appendix II, is a function of core build-up, and a critical ratio has been found to exist for toroidal cores of a given material. Core build-up is defined as the ratio of outside-to-inside diameter of the core. If core build-up exceeds the critical ratio, then μ_d decreases exponentially as build-up increases, but μ_d remains essentially constant for core build-up ratios less than the critical ratio. The critical ratio for Hipernik V core material is 1.40 to 1 and is somewhat smaller for the more sensitive core materials.[9] The cores used in the experimental work described herein have build-up ratios considerably smaller than the critical ratio.

FREQUENCY

If the minimum area has been selected to meet the volume requirements for input power and response time, the supply voltage for the core is a function of the number of gate winding turns and the frequency. If a large amount of the available winding area is to be used for control copper, the supply voltage and power output can only be increased by increasing the frequency. This is the real advantage to be gained by the higher carrier frequencies. Of course, the upper limit on the voltage for many applications is imposed by the rectifier ratings. An upper limit on frequency also exists because of winding capacitance and exciting currents. This upper limit on frequency, however, is well above the audio-frequency range. If a transistor power supply is used, a reasonable frequency for efficient operation is approximately 10,000 cps.

An important frequency effect is that, as frequency increases, eddy currents in the core decrease the effective a-c permeability. This decrease has been quantitatively evaluated by testing cores in a half-wave self-saturating circuit. The supply voltage was sinusoidal, and transfer curves of amplifier operation were obtained at a number of frequencies. Values of the effective a-c permeability μ_d were evaluated from the curves and these values are plotted as curve A in Fig. 2. Curve B is an approximate expression for the variations of μ_d with frequency. This approximate equation is $\mu_d = C/A + \sqrt{f}$ where C/A is the d-c permeability. Over a frequency range from approximately 3 to 10 kc, μ_d may be approximated by $\mu_d = C/\sqrt{f}$. The design equations previously discussed are modified by this frequency effect and the modified equations are given in Appendix III.

In addition to decreasing μ_d, high frequencies also accentuate capacitive effects, and increase exciting current and the saturated impedance of the core. These effects, which are difficult to evaluate quantitatively, prevent accurate prediction of amplifier characteristics by means of the design equations developed in this paper. Despite the lack of accuracy, however, the equations show that it is advantageous to decrease input power and time constant by decreasing the core size and that the figure of merit can be increased by operating the amplifier at a high frequency. The design equations have been followed in the development of the amplifier described in the following section of the paper.

Operational Amplifier Design Considerations

The operational amplifier described herein has been developed to function as a preamplifier in a 60-cps servo system. The following set of specifications were used as a working guide to develop the amplifier.

1. Time constant, 0.005 second or less.
2. Maximum available control power, 3×10^{-6} watt.
3. Maximum power output, 5 milliwatts.
4. Supply voltage available, d-c and 60 cps.
5. Supply voltage regulation, ±10 per cent.
6. Ambient temperature, −55 to +71 degrees centigrade (C).
7. Input resistance, not less than 2,000 ohms.
8. Maximum gain change under adverse environmental and supply voltage conditions is not to exceed ±10 per cent of nominal gain.
9. Zero drift level is not to exceed 10^{-6} watt.
10. Amplifier must provide reversible polarity output for reversible polarity input signals.

To meet these requirements, Hymu *80* core material and silicon diodes were chosen as basic components for the operational amplifiers. Hymu *80* core material, which was chosen because of its sensitivity, has the following properties:

1. Saturation flux density is 6,800 gausses.
2. ΔH for full control over the linear

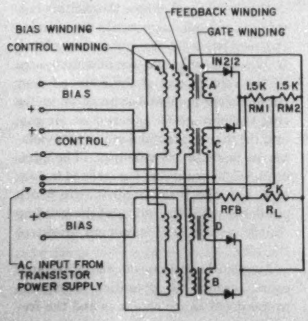


Fig. 3 (left). Push-pull magnetic amplifier

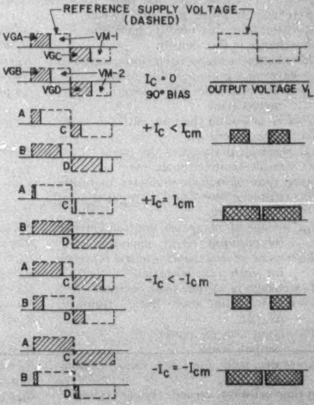


Fig. 4 (right). Voltage wave forms for square-wave supply

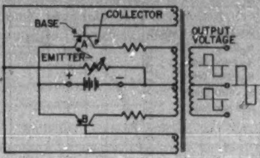


Fig. 5 (left). Transistor power supply

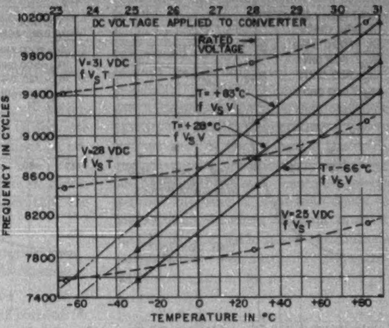


Fig. 6 (right). Characteristics of transistor power supply

region as determined by half-wave tests is 0.032 oersted at 10 kc.

3. Effective a-c permeability μ_d is 1.42×10^4 at 10 kc.

4. Strip thickness is 1 mil and strip width is 1/8 inch.

Silicon diodes were chosen for rectifiers to reduce capacitive effects and to decrease reverse leakage. The inverse voltage ratings of the type of rectifiers available was 64 volts. Because of the push-pull circuit characteristics, this limited the supply voltage to 32 volts.

The core size used in the amplifier was chosen by matching as closely as possible the core volume to the product of maximum available input power and specified time constant. The core dimensions were chosen by limiting the minimum number of turns of core material to six, which made it feasible to use a 3/4 inch by 1-inch core box size. The core area with six wraps of material is 0.0048 cm.^{2} Complete information on turns is given at the end of this section.

The rectifiers used with this amplifier have a maximum current rating of 20 milliamperes and the values of the mixing resistors and load resistor were selected not to maximize gain but to limit the load current to the current rating of the rectifiers. Negative current feedback is used in the operational amplifier to improve linearity, stability, and response time. The feedback voltage, which is obtained from the load resistor, is applied to the cores through 500-turn feedback windings. A feedback resistor is used to adjust the amount of current feedback. A forcing resistor is connected in series with the control windings to increase the input resistance of the amplifier.

The circuitry used for the final design of the amplifier is the conventional self-saturating circuit with mixing resistors. The supply voltage, however, is a square-wave voltage obtained from the transistor power supply. The diagram of the magnetic amplifier is shown in Fig. 3. The cores are wound with four windings: gate, control, bias, and feedback. The self-saturating rectifiers are silicon diodes that have an inverse voltage rating of 64 volts. The cores are matched for gain and the separate bias arrangement corrects for any mismatch in bias requirements between each pair of cores. The load resistor is 2,000 ohms and the mixing resistors are 1,500 ohms each.

Circuit operation is the same as for a conventional push-pull circuit. The output voltage wave forms differ somewhat because the supply voltage is a square wave. Wave forms of the voltages appearing across the mixing resistors and the output are shown in Fig. 4 for various values and polarities of the control current. The cores are biased at 90 degrees firing angles.

The amplifier cores had the following specification: Hymu *80* material; tape-wound toroid configuration; 0.75-inch inside diameter; 0.125-inch tape width; and 0.001-inch tape thickness. The outside diameter for core *A* was 0.875 inch, and for core *B* 0.765 inch.

There were four windings:

1. Load winding: for core *A*—300 turns, no. 30 wire; for core *B*—2,500 turns, no. 36 wire.
2. Control winding: for core *A*—2,000 turns, no. 36 wire; for core *B*—2,000 turns, no. 39 wire.
3. Bias winding: for core *A*—100 turns, no. 36 wire; for core *B*—100 turns, no. 39 wire.
4. Feedback winding: for core *A*—500 turns, no. 36 wire; for core *B*—500 turns, no. 39 wire.

The rectifiers were type *1N212* silicon diodes, one cell per leg. The supply voltage was 11.25 volts, 3,100 cycles for core *A*, 11.25 volts, 3,100 cycles for core *B*, and 23.50 volts, 6,175 cycles for core *B*. N_c^2/R_c was 5,450 for all cores. The operational amplifier was 30.00 volts, 8,700 cycles, with $N_c^2/R_c = 2{,}300$.

The power supply cores had the same specifications as the amplifier cores with the same outside diameter as core *A* and the following three windings:

1. Primary: 234 turns on each side of the center tap no. 29 wire.
2. Secondary: 250 turns on each side of the center tap no. 29 wire.
3. Bias: two windings of 42 turns each, no. 29 wire.

The transistors were Westinghouse type *2N71*, and the supply voltage was 28 volts d-c.

Transistor Power Supply

The circuit diagram for the power supply is shown in Fig. 5. It consists of two junction-type transistors and a transformer using a square-loop core material. The transistors act as controlled switches in a flip-flop type of operation. The d-c supply voltage is alternately applied to each half of the primary winding and produces an alternating square wave of voltage at the output of the secondary winding. The frequency is directly proportional to the average d-c supply voltage over a voltage range of approximately 20 to 1. Bias or "base" current is obtained from the d-c supply and is adjusted so that the transistors are well saturated when they are conducting. The output voltage is applied to the push-pull center-tapped magnetic amplifier circuit.

The square-wave oscillator makes effective use of the power ratings of the transistors but certain power supply limitations exist. Careful design is necessary to minimize exciting currents and the transistor ratings must be matched to extremes of ambient temperature. If the power supply is to supply only milliwatts of power, the core size must be kept low and a large number of turns is needed on the core. Otherwise, the exciting current would be larger than that drawn by the load. The internal power dissipation of the transistors is low because the transistors are operating as on-off switches in the power supply circuit. Consequently, they are capable of supplying power at a level from 20 to 50 times their

l, 74
1
55

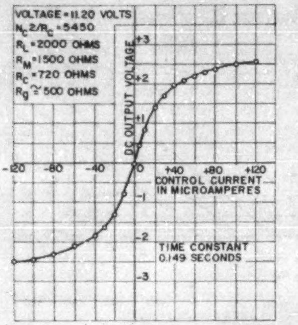


Fig. 7. Cores A operating in amplifier at 3,100 cycles

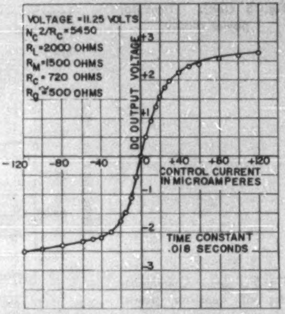


Fig. 8. Cores B operating in amplifier at 3,100 cycles

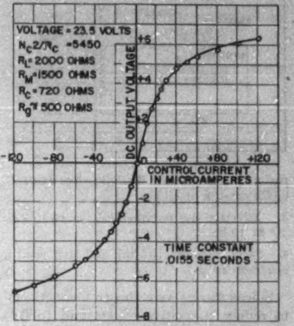


Fig. 9. Cores B operating in amplifier at 6,175 cycles

normal rating. At high ambient temperatures, however, their ability to deliver power is greatly lessened by thermal runaway effects and considerable derating is necessary.

The transistor power supply for the operational amplifier was first built and tested with Westinghouse type-*2N74* transistors, which have a 200-milliwatt dissipation rating at room temperature. With these transistors, the power supply was found to be capable of delivering 1 watt of power over an ambient temperature range of −65 to +28 C. At higher ambient temperatures, however, the power output was greatly diminished. To ensure satisfactory dissipation, the *2N74* transistors were replaced with Westinghouse type-*2N71* transistors which have the same junction area as the *2N74* but which have radiating fins to increase the permissible heat dissipation. The power supply using these transistors operated satisfactorily at ambient temperatures as high as 83 C. From the foregoing discussion it can be seen that, if high operating temperatures are to be encountered and if high efficiency is desired, careful considerations must be given to the design of the oscillator circuit and to the choice of transistors.

STABILITY

Stability of the transistor power supply was determined by operating the power supply and operational amplifier together over an ambient temperature range from −66 to +83 C with ±10-per-cent voltage variations of the d-c supply voltage. Curves showing plots of frequency versus voltage at constant temperature, and frequency versus ambient temperature with constant voltage are shown in Fig. 6. The results of the tests show that frequency varies linearly with voltage at a constant ambient temperature, but the frequency changes +4 per cent and −3 per cent over the ambient temperature range with rated voltage (28 volts d-c) applied.

Experimentation

The role of the modified equations in the development of the low-level operational amplifier is to indicate the effects on input power, response time, and figure of merit of the magnetic amplifier as core dimensions and frequency are varied. An experimental check is made by comparing computed values of input power, time constant, and figure of merit with measured values for amplifier operation with cores of different iron volumes at two frequencies. The cores are made of the same core material (1/8-inch by 1-mil Hymu *80*) and have the same inside diameter. The *A* cores have a cross-sectional iron area of 0.04 cm² and a mean magnetic path length of 6.50 cm. The *B* cores are specially constructed cores with six turns of the core material; they have a cross-sectional area of 0.00484 cm² and a mean magnetic path length of 6.04 cm.

The cores are operated in the circuit of Fig. 3 with the same supply voltage, rectifiers, load and mixing resistors, and N_c^2/R_c ratio. The turns in the gate windings of the *B* cores are adjusted so that they operate at the same voltage and frequency as the *A* cores. Both *A* and *B* cores are operated at the same frequency to check the equations for core variation effects, and the *B* cores are operated at a higher frequency to check the equations for the effects of frequency variation. The values of R_g are determined from the full-wave transfer curves for each section of the push-pull amplifier operating with infinite load resistance. The values of μ_g

Table I. Comparison of Measured and Computed Values

Core	Frequency, Cycles	Input Power, Microwatts		Time Constant, Seconds		Figure of Merit $\times 10^3$ per Second	
		Measured	Computed	Measured	Computed	Measured	Computed
A	3,100	0.122	0.118	0.140	0.215	0.308	0.20
B	3,100	0.122	0.102	0.018	0.027	3.08	1.80
B	6,175	0.184	0.203	0.0155	0.020	11.80	5.07
*B**	8,700	0.68	0.675	0.0055	0.007	20.6	8.64

* These measurements are for *B* cores operating in the low-level operational magnetic amplifier without feedback but with a 1,000-ohm forcing resistor connected in series with the control circuit.

Table II. Comparison of Measured and Computed Ratios

Compared	Linear Range of Output	Input Power, $P_{in}A/P_{in}B$		Time Constant, TA/TB		Figure of Merit, MA/MB	
		Measured	Computed	Measured	Computed	Measured	Computed
A/B	(+1)−0−(−1)	1.30	1.16	8.27	7.70	0.102	0.111

are obtained from curve *B* of Fig. 2 and time constants are determined from amplifier frequency response measurements.

EXPERIMENTAL CHECKS OF DESIGN EQUATIONS

Amplifier transfer curves of output voltage versus control current are plotted in Figs. 7 and 8 for amplifier operation with *A* and *B* cores at 3,100 cycles. The curve for amplifier operation at 6,175 cycles with the *B* cores is shown in Fig. 9. Measured values for input power and power gain are determined for each amplifier operation from the characteristic transfer curves and, together with the measured time constants for each operation, they determine the measured values of input power, time constant, and figure of merit. The modified equations are used to compute the values for these quantities for each amplifier operation. A comparison of measured and computed values for input power, time constant, and figure of merit is shown in Table I. The quantities are listed for amplifier operation over the linear region of the curve.

Table II shows a comparison of measured and computed ratios of input power, time constant, and figure of merit for *A* and *B* cores operating in the amplifier at 3,100 cycles. The trends in the variations of input power, time constant, and figure of merit predicted by the equations for variations of core size and frequency are verified by the experimental results. Good correlation is obtained between computed and measured ratios when the equations are used to predict changes in the quantities caused by core variation, but only a fair correlation is obtained when frequency is variable. As previously stated, the multitude of effects on amplifier characteristics caused by high-frequency operation is difficult to evaluate and prevents accurate quantitative predictions by means of the equations.

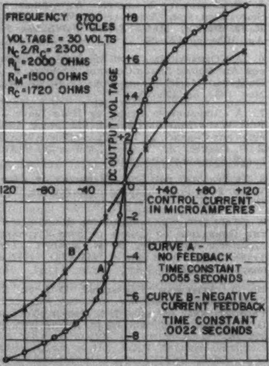


Fig. 10. Low-level operational magnetic amplifier characteristic curves

The experimental results, however, verify that it is advantageous to decrease the input power and time constant by decreasing the volume of the core. The power output, hence the figure of merit, can then be increased by operating the amplifier at a high frequency.

EXPERIMENTAL CHARACTERISTICS OF OPERATIONAL MAGNETIC AMPLIFIER

The characteristic transfer curves of output voltage versus control current for the operational amplifier are shown in Fig. 10. Curve *A* is obtained for amplifier operation without feedback, and curve *B* is obtained for amplifier operation with negative current feedback applied. Amplifier characteristics for the two modes of amplifier operation are shown in Table III. The characteristics pertain to a change in output voltage of 4 volts.

Stability is evaluated by measuring the change in gain and the amount of control current necessary to restore a balanced amplifier for zero output when the amplifier is operating with ±10-per-cent changes in voltage and frequency over an ambient temperature range from −66 to +83 C. The measurements are made for the amplifier only operating over the ambient temperature range, and for both amplifier and transistor power supplies operating at the same ambient over the range.

The amplifier stability measurements are:

Changes in gain, ±6.5 per cent.

Zero drift (referred to control current), 2.5 micro amperes.

Zero drift (referred to control power), 1×10^{-8} watts referred to input power.

The amplifier and power supply stability measurements are:

Changes in gain, ±10 per cent.

Zero drift, 4.0 microamperes total.

Zero drift, 2.75×10^{-8} watts referred to input power.

Conclusion

Design equations have been developed which show that both the input power and the long-time delays of high power gain magnetic amplifiers can be reduced by decreasing the core size and that power amplification and figure of merit can be increased by operating the amplifier at high frequencies.

Based on this concept, a stable magnetic amplifier has been developed that has a figure of merit of approximately 500,000 per second, a time delay of 2.2 milliseconds, and a drift level less than 0.03 microwatt. The amplifier has a bandwidth considerably greater than magnetic operational amplifiers operating at power frequencies. It can be used at these frequencies as an operational amplifier for fast response feedback control systems without imposing its own time delay on the system. In block diagram form, it may be represented in many systems as a constant gain with negligible time constant.

Appendix I. Basic Equations

Design equations presented in this paper are based upon certain basic equations developed in previous investigations.[7,8,9] The previous work and the further development in this paper are based upon the following simplifying assumptions.

1. The change in preset core flux level is proportional to the change in net average mmf acting on the core during the flux-setting half-cycle.
2. Circuit parameters are constant.
3. Saturated impedance of the reactor is pure resistance with a constant value over the range of operation.
4. Exciting currents are neglected.
5. Rectifiers are ideal and have infinite reverse resistance.
6. Forward rectifier resistance is constant over the operating range.
7. Effect of bias circuitry is negligible.
8. Supply voltage saturates the cores.
9. The cores are biased for 90 degrees' firing angles with zero control signal.
10. The ratio of the number of turns squared per ohm that can be wound into a given area is proportional to the area.
11. The effective a-c permeability of a given core material is constant for toroidal cores whose outside-diameter-to-inside-diameter ratio is less than 1.40.

The design equations developed are all based upon the assumption that the change in preset core flux level of a core during its flux-setting half-cycle is porportional to the change in net average control ampere-turns during the flux-setting half-cycle. Thus

$$\Delta\phi_m\times10^{-8}=KN_c\Delta I_{cm} \qquad (1)$$

K is the slope of the locus of reset points as shown in Fig. 1(A).

Table III. Amplifier Characteristics

	Curve A, No Feedback	Curve B, Negative Feedback
ΔP_{out}, milliwatts	8.0	8.0
ΔP_{in}, microwatts	0.68	5.6
K_p	11,800	1,430
T, seconds	0.0055	0.0022
M per second	2.06×10^6	6.5×10^5

The maximum control current needed to control a core in a push-pull circuit will vary with the biasing point and circuit parameters as explained by Hauge.[8] However, to simplify the development of the design equations, the maximum control ampere-turns needed for control in a push-pull circuit $N_c\Delta I_{cm}'$, as shown in Fig. 1(A), is assumed to be one-half the maximum control ampere-turns $N_c\Delta I_{cm}$ needed for complete control of a single core.

Thus

$$\Delta I_{cm}'=\frac{\Delta\phi_m\times 10^{-8}}{2KN_c}$$

and the equation for maximum input power can be written directly as

$$(\Delta P_{in})_m=(\Delta I_{cm}')^2R_c=\frac{1}{4K^2}\,\frac{\Delta\phi^2}{N_c^2/R_c}\times 10^{-8}\text{ watt} \quad (2)$$

The time constant in seconds from reference 9 is

$$T=2KN_c^2/R_c \quad (3)$$

The product of input power and time constant is

$$\Delta P_{inm}T={}^1/_2K\Delta\phi^2\times 10^{-8}\text{ watt-second} \quad (4)$$

The current gain from reference 9 is

$$K_I=\frac{2\omega KN_gN_c}{\pi R'}$$

The figure of merit on a per-second basis is

$$M=(K_I)^2\frac{R_L}{R_c}\,\frac{1}{T}$$

By substitution

$$M=8KR_L\left(\frac{N_g}{R'}f\right)^2\text{ per second} \quad (5)$$

where

$$R'=2R_g+R_L+R_LR_g/R_m$$

Appendix II. Design Equations

The basic equations may be modified to include core dimensions and properties by evaluating the constant K of equation 1 of the basic equations in terms of core properties.

From equation 1

$$\Delta\phi_m\times 10^{-8}=KN_c\Delta I_{cm}$$

The constant K is a measure of core gain and can be evaluated from half-wave circuit tests in terms of the magnetic properties of the core.

By definition

$$\Delta\phi=\Delta BA_c$$

$$N_c\Delta I_c=\frac{\Delta Hl_m}{0.4\pi}$$

$$\mu_d=\frac{\Delta B}{\Delta H}=\frac{2B_m}{\Delta H_m}$$

μ_d is thus defined as the effective a-c permeability of the core and is the slope of the locus shown in Fig. 1(B).

By substitution

$$K=0.4\pi\frac{A_c}{l_m}\,\frac{\Delta B}{\Delta H}\times 10^{-8}=0.4\pi\frac{A_c}{l_m}\mu_d\times 10^{-8} \quad (6)$$

From the basic equations the control current for 90 degrees' bias point is as follows

$$\Delta I_{cm}'=\frac{1}{2K}\,\frac{\Delta\phi}{N_c}\times 10^{-8}$$

Substituting and letting $\Delta\phi=\Delta BA_c$

$$\Delta I_{cm}'=\frac{\Delta H_ml_m}{0.8\pi N_c}$$

The input power for control over the linear region

$$\Delta P_{in}=\frac{\Delta H_m^2l_m^2}{0.64\pi^2}\,\frac{R_c}{N_c^2}$$

By letting

$$N_c^2/R_c=\beta l_m^2$$

$$\Delta P_{in}=\frac{\Delta H^2}{0.64\pi^2\beta}=K_c\frac{\Delta H^2}{\beta}\text{ watts} \quad (7)$$

Substituting equation 6 into equation 3, the time constants in seconds are

$$T=(2)(0.4\pi)\frac{A_c}{l_m}\,\frac{\Delta B}{\Delta H}\,\frac{N_c^2}{R_c}\times 10^{-8}\text{ second}$$

$$T=0.8\pi\beta\frac{\Delta B}{\Delta H}A_cl_m\times 10^{-8}\text{ second}$$

$$A_cl_m=V$$

$$T=0.8\pi\beta\frac{\Delta B}{\Delta H}V\times 10^{-8}$$

$$T=K_{01}'\beta\frac{\Delta B}{\Delta H}V=K_{01}'\beta\mu_dV\text{ seconds} \quad (8)$$

and the product of input power and time constant is

$$\Delta P_{in}T=\frac{\Delta H^2\mu_dV\times 10^{-8}}{0.8\pi}$$

Multiplying equations 7 and 8

$$\Delta P_{in}T=K_{01}'K_0\Delta H\Delta BV \quad (9)$$

Substituting equation 6 into equation 5 and letting $(N_g^2/R_w)=\alpha l_m^2$, gives the figure of merit

$$M=\frac{R_wR_L}{R'^2}f^2\alpha l_mA_c\frac{\Delta B}{\Delta H}\times 10^{-7}\text{ per second}$$

$$M=K_{03}\frac{R_wR_L}{R'^2}\alpha\frac{\Delta B}{\Delta H}f^2V=K_{03}\frac{R_wR_L}{R'^2}\alpha\mu_df^2V \quad (10)$$

From assumption 3, $R_g=(R_w+R_F)$. For small values of winding resistance

$$\frac{R_wR_L}{R'^2}\cong\frac{R_L}{4R_mR_w}$$

Appendix III. Equations Modified to Show Frequency Effects

The quantitative decrease in the effective a-c permeability μ_d of the core as a function of frequency was evaluated by testing cores in a half-wave self-saturating circuit. Curve A of Fig. 2 is a plot of μ_d versus frequency determined from the tests, and curve B is an approximation of the variations of μ_d with frequency. The approximate equation of curve B is $\mu_d=C/A+\sqrt{f}$, where C/A is the d-c permeability of the core. Over a frequency range from approximately 3 to 10 kc, μ_d may be approximated by $\mu_d=C/\sqrt{f}$. The design equations are modified by this frequency effect and the modified design equations are

$$K=0.4\pi\frac{A_c}{l_m}\,\frac{C}{\sqrt{f}}\times 10^{-8} \quad (11)$$

$$T=K_{01}'\beta\frac{C}{\sqrt{f}}V\text{ seconds} \quad (12)$$

$$\Delta P_{in}T=\frac{C}{0.8\pi}\,\frac{\Delta H^2}{\sqrt{f}}V\times 10^{-8}\text{ watt-second} \quad (13)$$

$$M=K_{02}C\alpha\frac{R_wR_L}{R'^2}Vf^{3/2}\text{ per second} \quad (14)$$

References

1. Elimination of Asymmetry Zero-Drift Errors in Magnetic Servo Amplifiers, **W. A. Geyger.** *AIEE Transactions*, vol. 74, pt. I, May 1955, pp. 239–45.

2. An All-Magnetic Audio Amplifier System, **J. J. Suozzi, E. T. Hooper.** *Ibid.*, July 1955, pp 297–301.

3. Magnetic Frequency Multiplier, **L. J. Johnson, S. E. Rauch.** *Ibid.*, vol. 73, pt. I, Nov. 1954, pp. 448–52.

4. Analysis and Design of a Magnetic Frequency Multiplier, **Otto J. M. Smith, Jalal T. Salihi.** *Ibid.*, vol. 74, pt. I, Mar. 1955, pp. 99–106.

5. Operational Magnetic Amplifiers for Servomechanism Application, **W. F. Horton, R. O. Decker.** *AIEE Special Publication S-63*, Apr. 1954, pp. 123–33.

6. An Operational-Type Magnetic Amplifier for Air-Borne Servo-Control Systems, **R. M. Hubbard.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 425–30.

7. A Comparative Study of Half-Wave and Full-Wave Self-Saturating Magnetic Amplifiers for Control of Two-Phase Servo Motors, **R. O. Decker.** *M. S. Thesis*, University of Pittsburgh, Pittsburgh, Pa., 1954.

8. A Study of the Load Circuitry of a Full-Wave Push-Pull Magnetic Amplifier, **T. A. Hauge.** *Ibid.*

9. Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers, **R. W. Roberts.** *AIEE Transactions*, vol. 73, pt. I, 1954 (*Jan. 1955 section*), pp. 682–90.

10. A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage, **G. H. Royer.** *Ibid.*, vol. 75, pt. I, July 1955, pp. 322–26.

A Transducer-Type Frequency Meter

U. L. SMITH
NONMEMBER AIEE

INDICATING or recording frequency meters in the 1-per-cent commercial frequency class use various circuits to obtain their indication. In all these circuits the instrument mechanism forms part of the circuit. The cross-coil type, the moving-iron type and the vibrating-reed type all have mechanisms that are built specifically for these meters, and these mechanisms have very little other commercial uses.

A newly developed frequency transducer described here uses a circuit that converts frequency into d-c millivolts. Any standard d-c millivoltmeter, either indicating or recording, may be used to measure its output, therefore simplifying and broadening the scope of frequency meters.

Several advantages are provided by this development.

1. It permits the use of a standard d-c indicating or recording voltmeter by merely changing the instrument calibrating resistance.
2. Any size or shape of instrument may be used.
3. Ranges as narrow as 59 to 61 cycles can be built, but broad ranges such as 50 to 70 cycles are just as practical.
4. The transducer can be furnished for any basic frequency such as 50, 60, or 400 cycles.
5. Only one size and type of transducer is needed.
6. The scale is nearly uniform over the entire range.
7. It eliminates the necessity for tooling, building, and stocking of mechanisms that are used only for frequency meters.

The circuit used was chosen after investigating a number of possible arrangements. The requirements set up were:

1. That the circuit would permit the use of an existing type of mechanism and that the same type of mechanism and circuit be used for all sizes of instruments. This requirement can be appreciated for, at the present time, instruments with 90-degree scales usually use a cross-coil mechanism with some type of resonant circuit. Instruments with 250-degree scales cannot use cross-coil mechanisms and are forced to use single-element mechanisms with resonant circuits. Recording instruments use cross-coil, iron-cored mechanisms. Miniature instruments, if available at all, are usually restricted to the vibrating-reed type.
2. That the component parts of the circuit be the same for all basic frequency ranges. By this is meant that the resistors, condensers, and inductances be the same for a 58–60–62-cycle range as for a 50–60–70-cycle range and that the calibration be obtained by adjusting one or more of these components.
3. That it meet the standard requirements set up for electrical indicating instruments which are: accuracy, 3 per cent; voltage influence, 1/2 per cent for ±10-per-cent variation; temperature influence, 1 per cent for ±10 degrees Centigrade; and the burden not over 5 volt-amperes at the central point. In addition, there is a basic frequency or line frequency, and the meter reads the deviation from this frequency. It was felt that the accuracy at this point would be better if the circuit was null reading at this frequency.

Any circuit where the current varies with frequency can be used for measuring frequency. Not all these circuits are suitable or desirable for commercial frequency meters. A single circuit, using a condenser or inductance or a combination of the two, lacks sensitivity and has no zero current or balance point. This circuit requires the use of a suppressed zero instrument and can be used only on very wide frequency ranges. Resonant circuits were found to be necessary as it is not possible to obtain enough energy change to operate an instrument for a narrow range of 58 to 62 cycles. However, resonant circuits that have the resonant point at the line frequency or balance point require exceedingly good condensers, and these circuits are difficult to compensate for temperature. Bridge circuits using air-core inductances or condensers lack sensitivity. When the sensitivity is increased by using iron-cored reactors or iron-cored mutual inductances, the bridges are difficult to balance to zero reading because of the harmonics caused by the iron. In general, most bridge circuits either lack sensitivity or, with the mutual inductance bridge, they lack simple adjusting means for balancing. The circuit chosen is a form of bridge with the balancing arms both adjustable. This circuit is shown in Fig. 1.

The circuit uses two off-resonance parallel circuits. Each of these circuits contains resistance, inductance, and capacitance in series. The resistance is the same in each and is fixed. The condensers are 0.15 and 0.1 microfarad, and the inductances are 30 to 80 henrys adjustable. For a frequency range of 55-60-65 cycles, the inductance in circuit 1 is adjusted to give resonance at approximately 48 cycles. The inductance in circuit 2 is adjusted to give resonance at approximately 72 cycles. The current-frequency relation is shown in Fig. 2. At the 60-cycle point the currents through the two circuits are approximately equal. The currents I_1 and I_2 in these two circuits pass through the r resistors, the ring-type rectifier, and through the R_1 and R_2 resistors. The d-c millivoltmeter is connected across points X_1 and X_2 between the R_1 and R_2 resistors.

The four resistors marked r are equal, and their purpose is to limit the rectified current in one circuit from passing through the parallel path in the other circuit. The R_1 and R_2 resistors are the calibrated resistors, and the current-resistance drop across them determines the output of the transducer.

One of the requirements was that the circuit operate both indicating and recording instruments. This is impossible to do when using a-c mechanisms, as the burden of a-c indicating mechanisms is about 1 volt-ampere and for a recorder about 10 volt-amperes. The unit would be both too large in physical size and would require more than 5 volt-amperes to operate. However, as d-c mechanisms use only about 1/1,000 the energy of a-c mechanisms, the requirements can be met by adding a rectifier.

The rectifier used is selenium, and it is connected in series so as to give full-wave rectification in the circuit. Rectifier use in instruments is not new but, unfortunately, they are not always thought of as a reliable measuring device. Actually, when considered as current rectifiers, they change very little with time or temperature. They will remain nearly constant indefinitely unless damaged by excessive

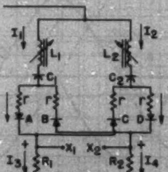


Fig. 1. Basic circuit showing direction of currents for first half-cycle

---

Paper 55-485, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 29, 1955; made available for printing June 1, 1955.

U. L. Smith is with the Westinghouse Electric Corporation, Newark, N. J.

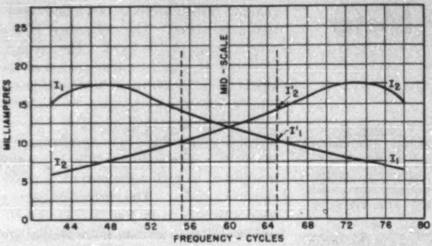


Fig. 2. Variation in branch currents with frequency change

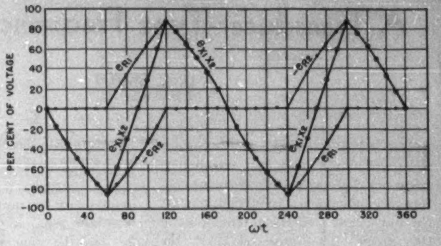


Fig. 5. Wave shape of voltages across R_1 and R_2 and X_1X_2 for 60 cycles

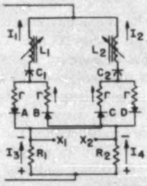


Fig. 3. Basic circuit showing direction of currents for second half-cycle

voltage or excessive high temperature. As current rectifiers, the only ageing is the change of back current or leakage current. In this application, the rectifier operates at 6 milliamperes at the central point and at 9 milliamperes at the end of the scale. The leakage current is in the order of 18 microamperes, which is less than 0.2 per cent of the total current. This means that normal ageing will not affect the accuracy. The a-c resistance through the rectifier does increase with time. It also decreases with increase in temperature. The resistance of the rectifier used is about 200 ohms, and the impedance of the circuit in which it is used is between 10,000 and 15,000 ohms. The rectifier then is less than 2 per cent of the total circuit. The effect of the temperature change is to change the impedance of the circuit but, in this case, is too small to be read. The ageing effect also changes the impedance and the sensitivity. This change is small in the overall circuit. It can be compensated for by selecting the correct values of *r* in Fig. 1. The current I_1 that passes through the *A* rectifier divides, the major part going through the R_1 resistor and a small part going through the rectifiers *C* and *D* and the resistors R_2. As rectifiers *C* and *D* age at the same rate as *A*, their aging effect will be to shunt more current through the resistor R_1 and thus compensate for the impedance change.

Circuit Explanation

The I_1 and I_2 currents are not in phase. However, to simplify the explanation of how the rectifiers work, they are for the time being assumed to be in phase.

Referring to Figs. 1 and 2, the currents at 60 cycles are equal. In circuit 1 on the first half-cycle, the I_3 current is downward in the circuit and passes through rectifier *A* and resistor R_1. (The I_3 current for the first half-cycle is I_1, and for the second half-cycle it is I_2. The I_4 current for the first half cycle is I_2 and for the second half cycle it is I_1.) The voltage drop across R_1 is I_3R_1, and the polarity can be considered as shown in Fig. 1. The current through circuit 2 for the first half-cycle is also downward and passes through rectifier *D* and resistor R_2. The voltage drop across R_2 is I_4R_2, and the polarity is the same as across R_1.

On the second half-cycle the current reverses and is upward in the circuits, as shown in Fig. 3. Current I_4 passes through resistor R_2 and rectifier *B*. The voltage across R_2 is I_4R_2, and the polarity is as shown in Fig. 3. The I_3 current passes upward through R_1 and rectifier *C*. The voltage across R_1 is I_3R_1.

Across the R_1 and R_2 resistors for the complete cycle there is an alternating voltage. However, at the X_1 and X_2 points, as the current $I_1 = I_2$ and the resistors R_1 and R_2 are equal, the voltage is at all times the same and has the same

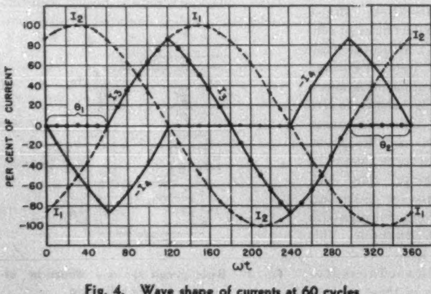


Fig. 4. Wave shape of currents at 60 cycles

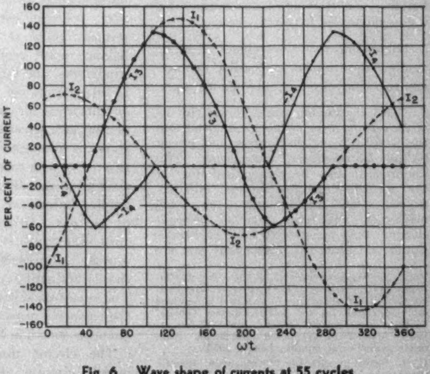


Fig. 6. Wave shape of currents at 55 cycles

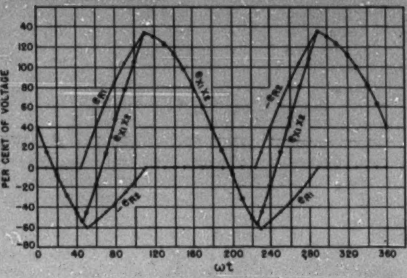


Fig. 7. Wave shape of voltages across R_1 and R_2 and X_1X_2 for 55 cycles

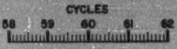


Fig. 9 (right). Scale marking showing uniformity of distribution

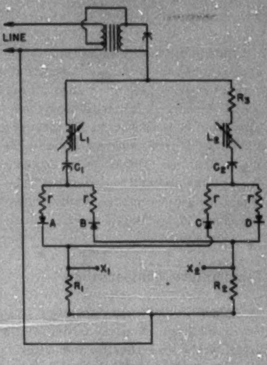


Fig. 10 (right). Complete frequency transducer circuit

polarity. There is then no direct voltage across these points and the millivoltmeter pointer is at mechanical zero. The instrument's mechanical zero will be the 60-cycle point on the dial and is normally at the center of the scale.

For a frequency of 65 cycles the currents in the two circuits are not equal, as shown on Fig. 2. The I_1 current has decreased to I_1', and the I_2 current has increased to I_2'. The direction and polarities are the same as for 60 cycles. The $I_2'R_1$ voltage is less than the $I_1'R_2$ voltage for the first half-cycle, and X_2 is more positive than X_1. On the next half-cycle the $I_1'R_2$ voltage is less than the $I_2'R_1$ voltage. The polarities, however, are reversed and X_2 is more positive than X_1, the same as for the first half-cycle. The voltage across X_1X_2 then is a full-wave rectified voltage with X_2 being positive with respect to X_1 and causes the pointer of the millivoltmeter to move to the 65-cycle point on the dial. For 55 cycles the reverse condition exists. The polarities change with X_1 being positive with respect to X_2.

Thus, it can be seen that a d-c millivoltmeter across X_1X_2 will read zero at the 60-cycle balance point and will read in one direction for an increase in frequency and in the other direction for a decrease in frequency. Its deflection is proportional to the difference between I_1 and I_2 or to the difference between the e_{R1} and e_{R2} voltages.

As pointed out in the foregoing, the currents I_1 and I_2 are not in phase. Circuit 1 is inductive, and the I_1 current lags the applied voltage. Circuit 2 is capacitive, and I_2 leads the applied voltage. At 60 cycles the angle of lead is equal to the angle of lag. The effect of these out-of-phase currents is to produce an alternating potential across X_1X_2.

Because of the rectifiers the current through R_1 and R_2 and the voltage across R_1 and R_2 and across X_1X_2 cannot be analyzed by rotating phasors. An expression can be given by breaking the complete cycle into sections. The graph of these values is given in Figs. 4 and 5. The voltages across R_1 and R_2 are alternating and have an odd-shaped wave. The voltage across X_1X_2 is alternating with a sawtooth wave. It is symmetrical about the X axis and is twice the fundamental frequency. The millivoltmeter across X_1X_2 will measure the average value of this wave, which is zero.

With a change of frequency, the value of current increases in one circuit and decreases in the other, and the phase angle decreases in the first and increases in the second. The instantaneous currents and voltages for the resistors at a frequency of 55 cycles are shown in Figs. 6 and 7. The voltage X_1X_2 is not symmetrical about the X axis but has a large positive value with a small negative value. A millivoltmeter connected with its negative terminal to the positive X_1 point and its positive terminal to the negative X_2 point will measure the average of this wave and will indicate 55 cycles on the dial.

Output Linearity

The rectified output voltage $e_{X_1X_2}$ is nearly proportional to the frequency. The value for this is shown in Fig. 8. It is symmetrical about the central frequency, and the instrument scale will be slightly condensed near the ends. The nearly

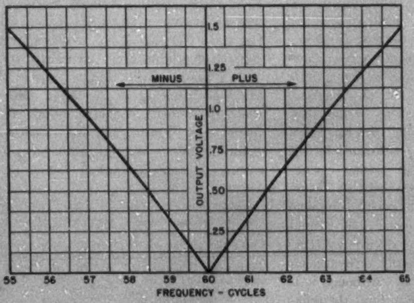


Fig. 8. Linearity of output with frequency change

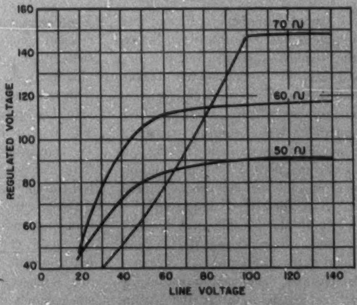


Fig. 11. Voltage regulator characteristics

linear scale distribution as it appears on the dial is shown in Fig. 9.

Sensitivity

Different frequency ranges for various applications are obtained by using millivoltmeters with different ranges. For a 50–60–70-range, a center zero d-c instrument is used requiring 0.5 milliampere and a series resistance of about 6,000 ohms. For other ranges, different values of resistance are used. Calibration is obtained by adjusting the millivoltmeter resistance to match the output of the transducer.

Temperature Compensation

Changes in ambient temperature have only a small effect on the output of the circuit. The basic effect of an increase in temperature would be to lower the output, but this is compensated for by replacing part of resistances R_1 and R_2 in Fig. 10 with positive coefficient resistance such as copper. In this way the output of X_1 and X_2 is kept constant for normal changes in temperature.

Voltage Regulator

In circuits of this kind, the d-c output is influenced by changes in line voltage except at the central point where the output value is zero. To make the entire scale independent of voltage change, a regulator is used. This is shown in the complete circuit in Fig. 10. The regulator is a saturated iron type and regulates between 100 and 130 volts. The primary and secondary voltage relations are shown in Fig. 11. As this is a saturation-type regulator, the secondary voltage changes with frequency. As shown in Fig. 11, at 50 cycles the secondary voltage is 90.5 and at 70 cycles 147.5 volts. The regulator will hold a constant secondary voltage within ±1/2 per cent for any frequency within its range, but at each frequency the voltage will be different. This means that, as the line frequency increases, a higher voltage is applied to the basic frequency circuit. The I_2 current in branch 2 will be greater than I_1. To make these currents nearer equal, the resistor R_3 is added in branch 2. The addition of the voltage regulator with its varying voltage with frequency characteristics changes the symmetry of the I_1 and I_2 relation, as shown in Fig. 2. The addition of the R_3 resistor reduced the I_2 current and flattened the curve. The difference between I_1 and I_2 below 60 cycles is now not the same as the difference between I_2 and I_1 above 60 cycles. These relations can be made near enough by retuning the circuit 2 to about 79 cycles. This can be seen from the scale distribution as shown in Figs. 8 and 9, which is the actual scale distribution for the complete circuit with voltage regulator.

Operating Characteristics

The burden of the complete transducer is 2.5 watts, 3.8 volt-amperes at 60 cycles. The voltage influence is within 1/2 per cent for 10-per-cent change in line voltage. The temperature error is less than 1/2 per cent for ±10-degree-centigrade change. A line voltage having a 10-per-cent 3rd harmonic will cause less than 1-per-cent error at the ends of the scale and very little error at the central frequency. The complete transducer is housed in a case 6 by 10 by $4^1/_8$ inches and it weighs 7 pounds.

Conclusion

This frequency transducer provides a new means of measuring frequency. It has the advantage that a basic unit can be calibrated with standard d-c instruments, either indicating or recording. Even for small panel boards, the frequency meter can be made to match, in size and shape, the rest of the instruments. It can be built with narrow frequency ranges for close frequency control or with wider frequency ranges for general use.

High-Frequency Operation of Self-Saturating Magnetic Amplifiers

H. WILLIAM COLLINS
ASSOCIATE MEMBER AIEE

THE supply frequency of magnetic amplifiers has been restricted largely to the readily available 60-cycle-per-second (cps) and 400-cps power sources. Recent developments in high-frequency power supplies which have ruggedness and reliability commensurate with magnetic amplifiers have made it practical to operate magnetic amplifiers at supply frequencies of over 400 cps. Several papers have described improved frequency multiplier techniques[1-3] which can be used to supply power for magnetic amplifiers in the audio range.

The recently developed transistor inverter[4] is possibly of greater future interest as a high-frequency power supply. The device is simple, small, rugged, and efficient, and operates the transistors in a switching mode to produce a power supply with characteristics commensurate with the inherent stability of a magnetic amplifier operated as a linear amplifier. Although at present the performance of the transistor inverter is limited by the frequency and output power characteristics of available transistors, future developments should greatly improve transistors in these respects. The vacuum tube equivalent of the circuit[5] will currently operate satisfactorily at frequencies in the megacycle range. Although it is not generally desirable to limit the reliability of magnetic-amplifier systems by using a vacuum tube, it is possible to use one twin triode tube to supply a number of magnetic amplifiers and still realize a reliability advantage compared to an all-vacuum-tube system.

The development of practical high-frequency supplies gives the magnetic-amplifier designer a new parameter, supply frequency, to vary arbitrarily over a wide range. The possibility exists of making magnetic amplifiers which are faster, smaller, and have greater power gain than attainable in the power frequency range. The primary purpose of this paper is to examine the design factors which determine the control power, response time, power gain, and the other characteristics of magnetic amplifiers operating at supply frequencies above the

---

Paper 55-465, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 30, 1955; made available for printing April 20, 1955.

H. WILLIAM COLLINS is with the Westinghouse Electric Corporation, Pittsburgh, Pa.

audio range, i.e., from about 20 kc to several megacycles.

Characteristics of High-Frequency Magnetic Amplifiers

The design considerations for high-frequency amplifiers are in many respects different from those of 60- or 400-cps amplifiers. The change in design characteristics and amplifier performance as supply frequency is increased can be most easily studied if it is assumed that an amplifier is to be designed to supply the same maximum load voltage across the same load resistance (constant maximum output power) at various supply frequencies. The most basic change in magnetic-amplifier design occurs because the volts per load turn supported by a given core are, of course, proportional to supply frequency. As a result, if a core cross-sectional area appropriate for the power frequency range is maintained, very few load turns are required to support a given supply voltage as supply frequencies approach the radio-frequency range.

High amplifier exciting currents can limit the degree to which it is practical to reduce the number of load turns in a high-frequency amplifier. The high core losses, combined with a small number of load turns, will increase the amplifier cutoff point and decrease the linear control region. Consequently, it is desirable to support a given voltage at high frequencies by reducing the iron cross section and keeping the number of load turns reasonably high. The control power and response time of a magnetic amplifier are also a function of the relationship between the core area and the number of turns selected to support a given voltage. It will be shown that it is also desirable to minimize the core area in order to keep the control power and response of the amplifier at a minimum.

L. 74
1
55

Nomenclature

A_b = cross-sectional area of core bobbin, square centimeters (cm)
A_p = amplifier power gain
b = height of core bobbin, cm
B_m = saturation flux density, gausses
ΔE_c = average control voltage increment
ΔE_L = average induced load winding voltage increment
E_s = a-c supply voltage, rms volts
f = amplifier supply frequency, cps
ΔI_c = control current increment required for full control of linear amplifier, see Fig. 1, amperes
I_L = average load current at maximum amplifier output, amperes
l_m = mean circumference of core, cm
L_s = saturated (air core) inductance of load winding, henrys
m = figure of merit in terms of power gain per cycle of response time T
M = figure of merit in terms of power gain per second of response time
N_c = number of turns in control winding
N_L = number of turns in load winding
P_c = control power required for full control of linear amplifier, Fig. 1, watts
P_L = average load power at maximum amplifier output, watts
r_1 = inside radius of core bobbin, cm
r_2 = outside radius of core bobbin, cm
R_c = control circuit resistance, ohms
R_L = load circuit resistance, ohms
T = response time in cycles of the supply frequency for 63 per cent output change with a step function input
X_s = saturated (air core) reactance in ohms of load winding at amplifier supply frequency
$\theta = \arctan^{-1} X_s/R_L$
μ_d = differential permeability. The ratio of reset flux density change Δ_B to control oersteds Δ_H occurring in the linear region when a magnetic material is operated in the mode of a simple, half-wave, self-saturating magnetic amplifier with high control circuit resistance.
μ_r = relative permeability

Control Power and Response Time

In the following derivation it is assumed that a self-saturating magnetic amplifier has the idealized transfer curve shown in Fig. 1, and that the slope of the transfer curve is the same as the curve that results when the core material is operated in an equivalent half-wave, self-saturating circuit of high control circuit resistance. The quantity ΔI_c is the control current required to drive the amplifier over the linear region from cutoff to full output.

The power for full control of the amplifier is

$$P_c = (\Delta I_c)^2 R_c \tag{1}$$

It has been shown previously[6] that the time constant of a self-saturating circuit expressed in cycles of the supply frequency is

$$T = {}^1/_2 \; \frac{\Delta E_L}{\Delta E_c} \; \frac{N_c}{N_L} \tag{2}$$

Selecting the average voltage increments to correspond to the linear region of Fig. 1

$$T = {}^1/_2 \; \frac{0.9E_s}{\Delta I_c R_c} \; \frac{N_c}{N_L} \tag{3}$$

The multiple of $P_c \times T$ for a given amplifier has been shown to be a constant over a wide value of control circuit resistances[7] and is a useful means of evaluating amplifier performance. Multiplying equations 1 and 3

$$P_c \times T = \frac{0.45 \Delta I_c E_s N_c}{N_L} \tag{4}$$

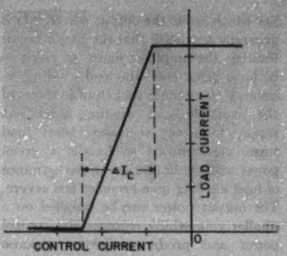


Fig. 1. Idealized magnetic amplifier transfer curve

The multiple of $\Delta I_c \times N_c$ is a function of the core material and is related to its magnetic characteristics[8] by the differential permeability. Selecting the ΔB, ΔH increments to correspond to the linear region of Fig. 1

$$\mu_d = \frac{\Delta B}{\Delta H} = \frac{2B_m}{\dfrac{0.4\pi N_c \Delta I_c}{l_m}} \tag{5}$$

$$\text{or } \Delta I_c = \frac{2B_m l_m}{0.4\pi N_c \mu_d} \tag{6}$$

The power for full control and response time can now be related and expressed in terms of core magnetic characteristics and other circuit constants. Substituting equation 6 in equation 4

$$P_c = 0.715 \frac{B_m l_m E_s}{T \mu_d N_L} \tag{7}$$

Equation 7 states that for a given response time and for a given core material the control power will be minimized if the supply voltage is supported with a large number of load turns and correspondingly small iron area. Summarizing the foregoing discussion, as supply frequency is increased, the alternative of supporting a given supply voltage with a reduced number of load turns or a reduced iron area exists. From the standpoint of minimizing exciting current, control power, and response time, it is desirable to work with the minimum core area which the load circuit considerations will allow.

Limitations on Magnetic Amplifier Performance

There is, of course, a limit to the number of load turns which can be placed on an amplifier core. In the power frequency range the maximum number of turns and minimum core area is set by the resistance of the load winding and the resultant internal impedance of the amplifier which limits the output power. It is generally accepted[7] that the major factor limiting the amplifier figure of merit *m* is, for a given core, the load winding resistance (sometimes spoken of in terms of the available load winding area). As supply frequency increases, fewer load turns are required to handle a given power and the limitation on performance of load winding area becomes less severe. The output power can be handled on a smaller core which requires less control power and produces shorter response times. Therefore, there is a possibility of achieving increased amplifier performance at high frequencies

Because of increased core losses, the differential permeability of magnetic materials (except possibly ferrites) decreases rather drastically with increasing frequency. Above the audio frequency range the differential permeability of even a 1/8-mil thickness magnetic tape is considerably lower than the differential permeabilities which can be obtained in the power frequency range. In contrast to the behavior of load-winding area, the limitation imposed on amplifier performance by the characteristics of core materials becomes more severe at high frequencies. High-frequency magnetic amplifier operation, as will be demonstrated, produces power gains per unit response time which are much higher than those attainable in the power frequency range. This increased performance is possible primarily because the beneficial effects which arise from the less severe limitations of load-winding area are greater than the detrimental effects which increased frequency has on the magnetic characteristics of the amplifier core materials.

Saturated Reactance

Most analyses of power frequency magnetic amplifiers are justified in assuming that the internal impedance of a saturated amplifier is purely resistive. It has been shown that as supply frequency is increased the limitation on performance by load-winding resistance is less severe. It must be recognized though that a completely saturated amplifier reactor retains an air core inductance. Air core inductances which are negligible at 400 cps can produce relatively large reactances at frequencies above the audio range. The reactance of a given winding is, of course, proportional to frequency, and the problem of load-winding reactance, unlike that of winding resistance, increases with increasing frequency. Intuitively it appears that at a high frequency the saturated reactance may be the major component of the amplifier internal impedance and therefore the major factor limiting the power output.

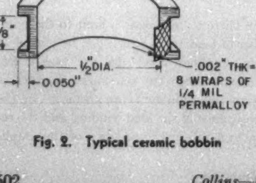


Fig. 2. Typical ceramic bobbin

To evaluate the relative importance of the resistive and the reactive components of the internal impedance of a high-frequency amplifier, it is necessary to examine the geometry of a typical core which might be used in such an amplifier. The most appropriate cores found for use in high-frequency amplifiers consist of a number of turns of ultrathin (1/8- to 1/2-mil) magnetic tape wound and annealed on ceramic bobbins[9] as shown in Fig. 2. As an example, data on a core appropriate for use in a 100-kc amplifier will be presented. To keep the exciting current and control power on such an amplifier small it is desirable to use several hundred load turns as shown in the following:

Core, 1/2-inch inside diameter×21/32-inch outside diameter×15/64-inch height ceramic bobbin.
Load turns, 400 turns no. 34 wire.
Load-winding resistance, 5.5 ohms.
Air core inductance of toroid, 112 μhenrys.
Air core reactance of toroid, 70 ohms.
Area consumed by load winding, 25 per cent of available winding area.
Output power of full-wave circuit, 2.0 watts at 25-volt supply.

At 100 kc the saturated reactance of the toroid is twelve times the resistance of the toroid, even though only a small part of the available winding area is consumed by the load winding. At higher frequencies the saturated reactance is large to an even greater degree, relative to the winding resistance. Likewise, as the number of turns is increased, the reactance increases with the number of turns squared and therefore reactance becomes the increasingly dominant component of the amplifier impedance.

These data indicate, and it has been experimentally verified, that for a wide range of output power at frequencies above the audio range the winding resistance is slight, relative to the load-winding saturated reactance. It is possible to generalize from this discussion that there is a fundamental difference between a high-frequency and a power frequency magnetic amplifier. In a power frequency amplifier the load-winding area (or winding resistance) is a major factor limiting amplifier performance; whereas at high frequencies the area required by the load winding is of little consequence, and the saturated reactance of the winding is the most basic limitation. With proper consideration of the nature of the internal impedance of a high-frequency amplifier, it is possible to determine the output power and derive expressions relating power ~~gain and~~ response time to the design variables.

Reactance-Limited Amplifier

The following analysis is for a bridge circuit shown in Fig. 3 with the idealized transfer curve of Fig. 1. The amplifier is operating at a supply frequency high enough to justify the following assumptions:

1. The load winding resistance is negligible, relative to the saturated reactance and load resistance of the amplifier.
2. The forward drop and back leakage of the rectifiers are negligible.
3. The circuit is not so highly inductive because of the saturated inductance that the ampere-turns required for full control of the amplifier are significantly altered from the purely resistive case.
4. The operation of the amplifier is not affected by winding capacitances.

In analyzing the effect of the saturated reactance on the output power of an amplifier, a simple means of relating the load current to the internal impedance of the amplifier is needed. For values of reactance which are small relative to the load resistance, it is a reasonably good approximation to assume that the load current at full output varies in the manner of a series linear circuit whose impedance is determined by the effect the reactance has on the phase angle θ. With values of saturated reactance smaller than the load resistance, the load current at saturation of the nonlinear circuit of Fig. 3 can there-

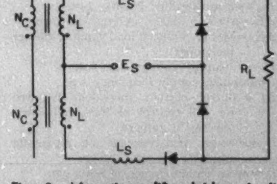


Fig. 3. Magnetic-amplifier bridge circuit with finite saturated reactance

Vol. 7
Pt. I
1955

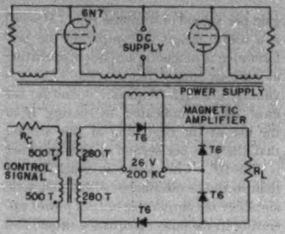


Fig. 4. Diagram of 200-kc magnetic amplifier and power supply

fore be expressed approximately as follows

$$I_L=0.9\frac{E_s}{R_L}\cos\theta=0.9\frac{E_s}{R_L}\left|\frac{R_L}{R_L+jX_s}\right| \tag{8}$$

At maximum output:

$$P_L=I_L^2R_L=\frac{0.81E_s^2\left(\frac{R_L}{R_L+jX_s}\right)^2}{R_L} \tag{9}$$

Note that output power is arbitrarily defined in terms of average current squared.

By definition

$$m=A_p\frac{1}{T}=\frac{P_L}{P_c}\frac{1}{T} \tag{10}$$

Substituting equations 7 and 9 in 10.

$$m=1.12\frac{E_s\mu_dN_L}{R_LB_ml_m}\left(\frac{1}{1+\frac{X_s^2}{R_L^2}}\right) \tag{11}$$

Examination of equation 11 shows that the figure of merit is a function of the X_s to R_L relationship. At very large values of R_L a small load is on the amplifier and the output power is small. At very small values of R_L the internal impedance of the amplifier (indicated by the term in parentheses) greatly decreases the output voltage and the output power is again small. Obviously there is an X_s-to-R_L relationship which yields a maximum figure of merit, and it is possible to determine this maximum by differentiation of m with respect to the X_s/R_L ratio.

Since there is an interrelationship between the saturated reactance X_s and the number of load turns N_L a substitution for N_L must be made before the differentiation is possible. The equation for the inductance of a single-layer wound toroid of rectangular cross section[10] provides a means of eliminating N_L as a variable by expressing it as a function of the saturated inductance L_s, which in turn can be expressed in terms of X_s.

$$L_s=\frac{4\times10^{-9}\mu_rN_L^2b(r_2-r_1)}{(r_2+r_1)} \tag{12}$$

$$X_s=2\pi f\frac{4\times10^{-9}\mu_rN_L^2b(r_2-r_1)}{(r_2+r_1)} \tag{13}$$

Note that

$\mu_r=1.0$ when amplifier is saturated
$A_b=b(r_2-r_1)$
$l_m/\pi=(r_2+r_1)$

Making the above substitutions and solving equation 13 for N_L

$$N_L=3{,}570\frac{X_s^{1/2}l_m^{1/2}}{A_b^{1/2}f^{1/2}} \tag{14}$$

N_L is expressed in terms of the other design factors and therefore can be eliminated as a variable from the figure of merit expression. Substituting equation 14 in equation 11

$$m=4{,}040\frac{E_s\mu_dX_s^{1/2}}{R_LB_ml_m^{1/2}A_b^{1/2}f^{1/2}}\left(\frac{1}{1+\frac{X_s^2}{R_L^2}}\right) \tag{15}$$

If the other quantities can be held constant while the saturated reactance X_s is varied, it is possible, by differentiation of equation 15, to determine the X_s/R_L ratio which will produce a maximum figure of merit. For a given core material, bobbin size, and frequency, the quantities μ_d, B_m, A_b, l_m, and f obviously can be held constant. The supply voltage E_s can be held constant if it is assumed that, as X_s is varied by varying N_L, the iron area is varied inversely with N_L, thus holding E_s constant. This condition assumes that the iron area can be arbitrarily varied without affecting the cross sectional area of the bobbin (A_b). Reference to Fig. 2 will demonstrate the justification of this assumption. For amplifiers of moderate output powers (up to 10.0 watts) a typical core for use above the audio frequency range has an iron buildup (r_2-r_1) of 0.004 inch or less. The smallest commercially available ceramic bobbin on which high-frequency tape is wound has a buildup of 0.050 inch. Therefore, the iron buildup can be arbitrarily varied from about 0.0005 to 0.030 inch (a 60-to-1 variation in iron area) and the area of the bobbin (A_b) will not be affected. This condition definitely does not exist in a power frequency amplifier, and the reader familiar with the problems of analyzing low-frequency amplifiers will realize that it vastly simplifies this analysis.

Since there is justification in holding the other variables constant, equation 15 can be rewritten in simplified form so that the dependency of the figure of merit on the X_s/R_L ratio may be studied.

$$\int\frac{dm}{dX_s}=C\frac{X_s^{1/2}}{R_L}\left(\frac{1}{1+\frac{X_s^2}{R_L^2}}\right) \tag{16}$$

Fig. 5. 200-kc magnetic amplifier and power supply

Differentiating and solving equation 16 for a maximum

$$X_s=1/\sqrt{3}R_L \text{ for maximum } m \tag{17}$$

Therefore, in supplying a given output power (given voltage across a given load resistance) the figure of merit will be maximized if the iron cross-sectional area is selected so that the load turns required to support the supply voltage have a saturated reactance which is $1/\sqrt{3}$ times the load resistance.

Experimental Amplifier

To check the characteristics of high-frequency magnetic amplifiers, a 200-kc amplifier has been tested. The amplifier was designed to supply approximately 2.0 watts to a 150-ohm resistive load and to be driven by a saturable core vacuum-tube type of inverter.[5] The power supply and amplifier circuits are shown in Fig. 4.

The amplifier was designed according to the relationships developed in the foregoing. The optimum saturated reactance for a 150-ohm load was determined from equation 17. A 1/2-inch ID ceramic bobbin was selected as being an appropriate core size for the output power and supply frequency. The supply voltage necessary to produce the required output power, and the number of load turns necessary to produce the desired saturated reactance were determined respectively from equations 8 and 14. If the supply frequency, supply voltage, number of load turns, and saturation flux density of the core material are known, it is possible to determine the iron cross sectional area required by the design. Equations 7, 11, and 15 indicate the performance which can be expected from the amplifier. The control circuit is designed in conventional fashion to produce the desired response time and the proper impedance match with the control source.

Fig. 5 shows the amplifier and power supply. In the following are summarized the design information and performance data for the amplifier.

Supply voltage E_s, 26 volts.
Load turns N_L, 280 turns no. 32 wire.
Control turns N_c, 500 turns no. 36 wire.
Ceramic bobbin size, 1/2×21/32×15/64-inch.
Core material, 7 wraps of 1/4 mil (0.00025-inch)×1/8-inch Hiperthin 4–79 Permalloy.
Differential permeability (μ_d) at 200 kc, 54,000.
Load-winding resistance, 2.5 ohms.
Load-winding saturated reactance, 70 ohms.
Rectifiers, transitron *T6* germanium diodes.
Control resistance for 1.0-cycle response, 8,000 ohms.
Measured gain per cycle response, *m*, 100.
Measured gain per second response, *M*, 20×10^6.

The effective differential permeability of the material was determined by operating the cores at 200 kc in a half-wave, self-saturating circuit which uses a vacuum-tube rectifier. Core losses can cause appreciable heating of the 1/4-mil core material when operating at 200 kc. Because of the extremely small core volumes (less than 0.002-inch build-up) the amount of heat generated is slight and the actual temperature rise of the metal is difficult to measure. Although the heat generated per unit volume is probably high, the rate of dissipation is rapid, and no detrimental effects on the amplifier behavior from core heating have been observed.

Small-winding capacitances like small-winding saturated inductances can affect the performance of high-frequency amplifiers. If an excessive number of load or control turns are placed on a high-frequency toroid, anomalous changes can occur in the amplifier control characteristic. Although winding capacitance is definitely a design consideration, careful design of the windings can usually eliminate its effects. The author has constructed amplifiers operating at supply frequencies as high as 2.0 megacycles which do not display detrimental capacitive effects.

A set of transfer curves for the 200-kc amplifier with various values of load resistance is shown in Fig. 6. The response time of the amplifier with various values of control resistance was measured by placing square waves from a pulse generator in the control circuit and observing the output transient with an oscilloscope. With a 150-ohm load this stage has a measured figure of merit of 100 per cycle. In comparison, a calculated figure of merit of 83 per cycle is obtained by substituting the design values in equation 11.

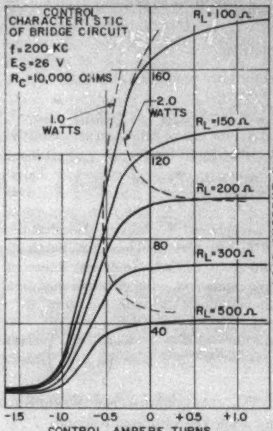


Fig. 6. Set of transfer curves for 200-kc magnetic amplifier bridge circuit

Fig. 7 is an oscillogram showing the response of the amplifier to a square wave input. The stage is operating with a delay of about 1 cycle (5.0 μ seconds) and a power gain of about 100.

Fig. 8 shows the output of the amplifier with 250-μ-second pulses injected into the control winding. A filter capacitor is used to suppress the 200-kc carrier. The control circuit resistance is selected so that the stage has a power gain of about 500 and a rise time for 63 per cent response of 25 μ seconds.

The figure of merit *M* on the basis of gain per second of response time ($M = mf$) is a useful means of comparing the performance of magnetic amplifiers operating at various supply frequencies. The advantages of high-frequency operation are most clearly indicated by noting from the list at the beginning of this section that the gain per second of this amplifier is 20×10^6. This value is much higher than would be obtainable with a magnetic amplifier of comparable output power operating in the power frequency range, and it makes possible superior performance with respect to both response time and power gain. One of the advantages of full-wave magnetic-amplifier circuitry is that it is possible to trade response time for power gain. For example, it is possible to give this amplifier a 1,000-cycle delay, achieve a power gain of 100,000 per stage, yet retain a response time (0.005 seconds), which is still less than the minimum half-cycle delay possible with a 60-cycle amplifier. There is naturally a limit to the N_c^2/R_c which can be placed on an amplifier, and therefore a limit to the delay and power gains possible, but delays of 1,000 cycles are reasonable. It is interesting to note that the performance of this amplifier is equivalent to that of a 60-cps amplifier with a figure of merit (gain per cycle) of 333,000. This value is at least two orders of magnitude greater than normally encountered with a 60-cps amplifier of a 2.0-watt output. As a result, high-frequency operation of magnetic amplifiers makes possible power gains and/or bandwidths unattainable in the power frequency range.

Conclusions

The operation of magnetic amplifiers at frequencies above the audio range is practical and advantageous. To minimize the control power, response time, and exciting current of high-frequency amplifiers the smallest iron cross-sectional area permitted by load circuit considerations should be used. As supply frequency increases, the limitation imposed on the performance of an amplifier of a given output power by the resistance of the

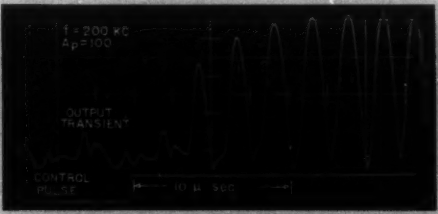


Fig. 7. Transient response of 200-kc magnetic amplifier bridge circuit

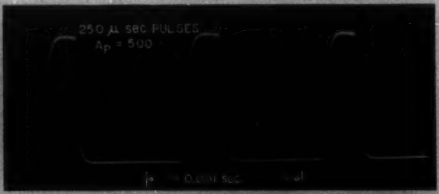


Fig. 8. 200-kc magnetic amplifier passing 250-μ-second pulses

load winding becomes less severe. At high frequencies the saturated reactance of the load winding is the major component of the amplifier internal impedance and therefore a major factor limiting amplifier performance. The figure of merit of a high-frequency amplifier can be expressed in terms of the design variables and is maximized when the load resistance equals $\sqrt{3}$ times the saturated reactance. Despite the relatively low differential permeabilities encountered in high-frequency operation, power gains per unit response time much higher than possible with power frequency amplifiers can be achieved because of the possibility of controlling a given output power with much smaller cores. As a result, magnetic amplifiers have been built which have power gains and bandwidths normally unattainable in the power frequency range.

References

1. Magnetic Frequency Multipliers, L. J. Johnson, S. E. Rauch. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 448–51.

2. Analysis and Design of a Magnetic Frequency Multiplier, Otto J. M. Smith, Jalal T. Salihi. *AIEE Transactions*, vol. 74, pt. I, Mar. 1955, pp. 99–106.

3. An All-Magnetic Audio Amplifier System, J. J. Suozzi, E. T. Hooper. *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 297–301.

4. A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage, G. H. Royer. *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 322–26.

5. A Variable Frequency Magnetic-Coupled Multivibrator, R. I. Van Allen. *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 356–61.

6. Response Time of Magnetic Amplifiers, E. L. Harder, W. F. Horton. *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1130–33.

7. The Figure of Merit of Magnetic Amplifiers, J. T. Carleton, W. F. Horton. *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 239–45.

8. Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers, R. W. Roberts. *AIEE Transactions*, vol. 73, pt. 1, Jan. 1954 (Jan. 1955 section), pp. 682–90.

9. Ultrathin Tapes of Magnetic Alloys with Rectangular Hysteresis Loops, Martin F. Littmann. *AIEE Transactions*, vol. 71, pt. I, July 1952, pp. 220–23.

10. Principles of Electric and Magnetic Fields (book). Warren B. Boast. Harper and Brothers, New York, N. Y., 1948, p. 249.

An Improved Detached-Contact-Type of Schematic Circuit Drawing

F. T. MEYER
MEMBER AIEE

l. 74
1
55

THE ever-increasing complexity of telephone and telegraph switching circuitry in the Bell System reached a stage, about 10 years ago, where many of the conventional type of schematic circuit drawings presented difficulties in training of personnel and maintenance. It then became necessary, for the more complex circuits, to prepare and maintain a separate series of drawings especially streamlined for training purposes and for aiding maintenance. However, the preparation of two sets of drawings for these cases and the necessity for keeping them in step as improvements and changes were added was costly. To meet this problem the new type of circuit drawing described in this paper was developed. It is designed for easy understanding and incorporates in a single package the basic circuit information required for manufacture, installation, training, maintenance, and engineering analysis.

General Description and Purpose of Schematic Circuit Drawing

Telephone and telegraph switching equipments, whether simple or complex, usually originate with a schematic circuit drawing which specifies the basic design from a circuit standpoint. It consists primarily of a layout of apparatus components such as relays, inductors, capacitors, and resistors, interconnected by a schematic circuit pattern.

Each such drawing becomes a permanent record of a specific design and is used for various purposes such as: 1. patent investigation, 2. engineering analysis, 3. the preparation of manufacturing drawings, and 4. shop testing. It is also used by the installation forces for testing, and for clearing troubles before cutover. Operating personnel use the drawing for both training and maintenance purposes.

Older Type of Schematic Circuit Drawing

The conventional older type of schematic circuit drawing in general use in the Bell System for many years includes the following:

1. Schematic circuit pattern: In general, apparatus components are symbolized in their entirety and shown as an integral part of the circuit pattern. For example, a relay core and windings and all its associated contacts are shown grouped together. The pattern arrangement and content of a schematic circuit figure is based on a predetermined physical association of apparatus. The pattern may be shown in one or more figures, as required, in order to identify different features or options. Fig. 1 shows a portion of the schematic circuit pattern of this type of drawing.

2. Circuit requirement tables, which give the testing and adjusting information for such items as relays, registers, and switches to insure proper operation of the circuit.

3. Cabling diagrams and terminal punching layouts, which show the proper cable and wire terminations, and the connection of the circuit into the system network.

4. Designations: Each apparatus component is assigned a functional designation (usually a lettered one) for readily identifying it in the circuit pattern. For relays and other elements listed in the circuit requirement tables, the functional designations are repeated in these tables.

Apparatus terminals, including relay spring terminals, are usually assigned the predetermined numbers which apply to the apparatus regardless of the drawing on which it is used.

Each apparatus component bears an identifying code number, for manufacturing and maintenance purposes. Code numbers are shown at the apparatus in the circuit pattern except in the case of relays where the code is shown only in the circuit requirement tables.

5. Supporting information, which includes working limits (voltage limits, conductor loop resistance, minimum insulation resistance, and ground potential differences), transmission test requirements for all circuits through which a talking path is established, optional wiring or apparatus or both, and certain tables and notes as required.

In addition, a circuit description is issued concurrently with each circuit drawing, and both are kept up to date. It includes such items as: 1. general description of circuit, 2. detailed description of circuit operation, 3. reference data, which includes (a) interpretation of functional

---

Paper 55-467, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 29, 1955; made available for printing May 6, 1955.

F. T. Meyer is with the Bell Telephone Laboratories, Inc., New York, N. Y.

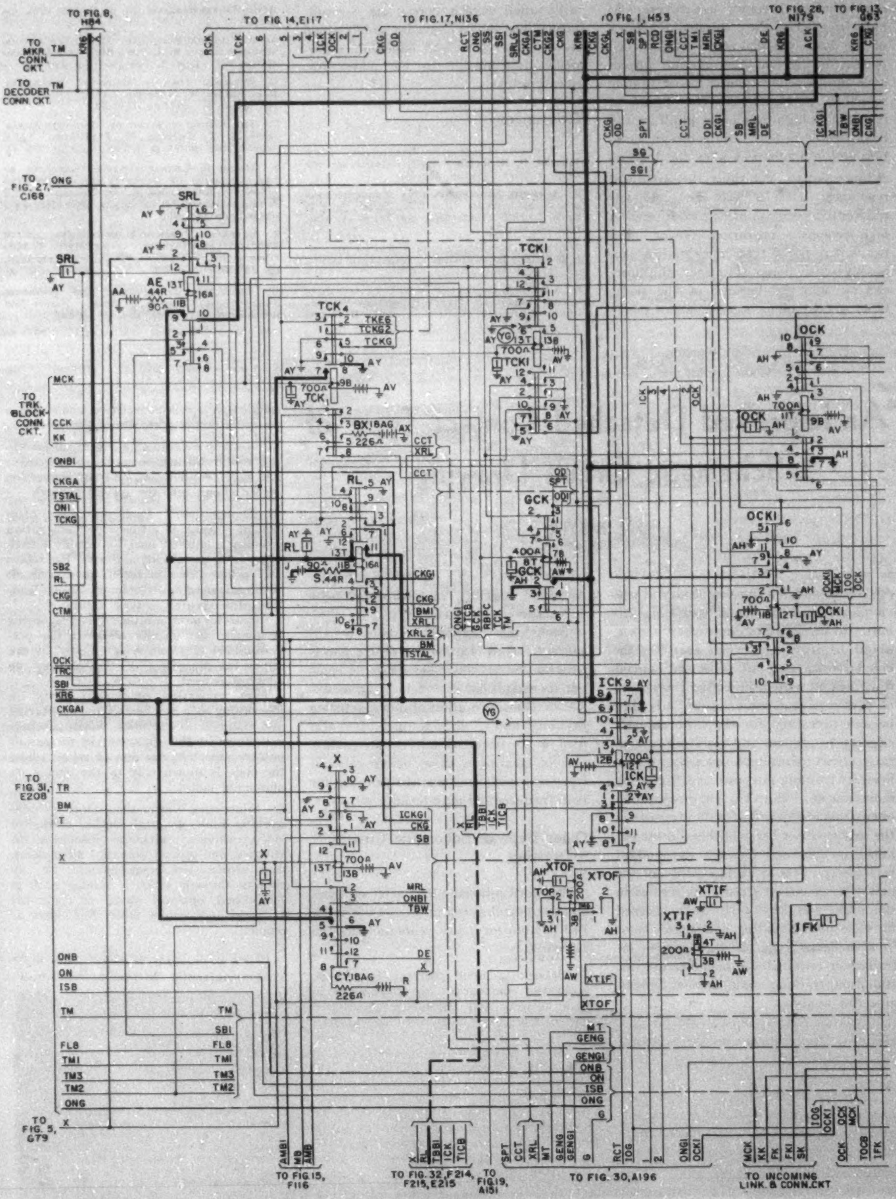

TO FIG 8, H84
TO FIG. 14, E117
TO FIG. 17, N136
TO FIG. 1, H53
TO FIG. 28, N179
TO FIG. 13, G83
TO MKR. CONN. CKT.
TO DECODER CONN. CKT.
TO FIG. 27, C168
TO TRK. BLOCK CONN. CKT.
TO FIG. 31, E208
TO FIG. 5, G79
SRL
AE
TCK
TCK1
BX
RL
S
GCK
OCK
OCK1
ICK
X
CY
XTOF
XTIF
IFK
TO FIG 15, F116
TO FIG. 32, F214, F215, E215
TO FIG. 19, A151
TO FIG. 30, A196
TO INCOMING LINK & CONN. CKT

Fig. 1. Circuit pattern of an attached-contact type of drawing

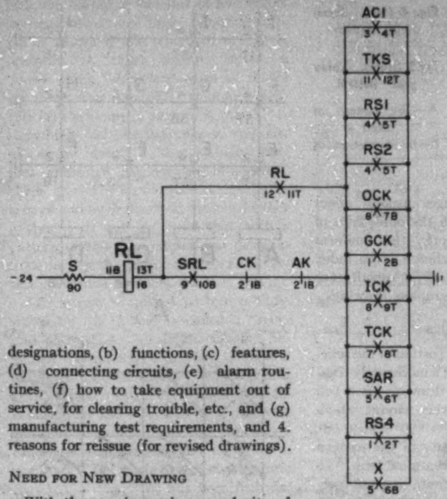
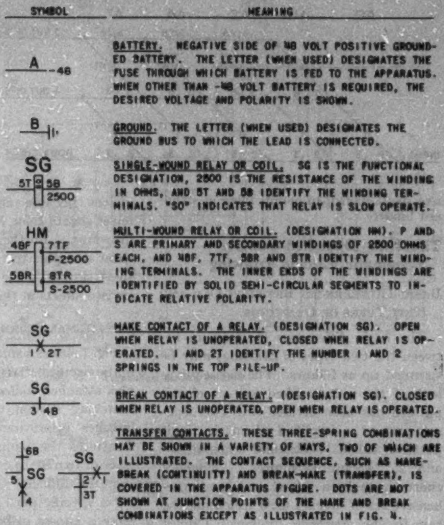


Fig. 2 (above). RL path in detached-contact form



BATTERY. NEGATIVE SIDE OF 48 VOLT POSITIVE GROUNDED BATTERY. THE LETTER (WHEN USED) DESIGNATES THE FUSE THROUGH WHICH BATTERY IS FED TO THE APPARATUS. WHEN OTHER THAN -48 VOLT BATTERY IS REQUIRED, THE DESIRED VOLTAGE AND POLARITY IS SHOWN.

GROUND. THE LETTER (WHEN USED) DESIGNATES THE GROUND BUS TO WHICH THE LEAD IS CONNECTED.

SINGLE-WOUND RELAY OR COIL. SG IS THE FUNCTIONAL DESIGNATION, 2500 IS THE RESISTANCE OF THE WINDING IN OHMS, AND 5T AND 5B IDENTIFY THE WINDING TERMINALS. "SO" INDICATES THAT RELAY IS SLOW OPERATE.

MULTI-WOUND RELAY OR COIL. (DESIGNATION HM). P AND S ARE PRIMARY AND SECONDARY WINDINGS OF 2500 OHMS EACH, AND 4BF, 7TF, 5BR AND 8TR IDENTIFY THE WINDING TERMINALS. THE INNER ENDS OF THE WINDINGS ARE IDENTIFIED BY SOLID SEMI-CIRCULAR SEGMENTS TO INDICATE RELATIVE POLARITY.

MAKE CONTACT OF A RELAY. (DESIGNATION SG). OPEN WHEN RELAY IS UNOPERATED, CLOSED WHEN RELAY IS OPERATED. 1 AND 2T IDENTIFY THE NUMBER 1 AND 2 SPRINGS IN THE TOP PILE-UP.

BREAK CONTACT OF A RELAY. (DESIGNATION SG). CLOSED WHEN RELAY IS UNOPERATED, OPEN WHEN RELAY IS OPERATED.

TRANSFER CONTACTS. THESE THREE-SPRING COMBINATIONS MAY BE SHOWN IN A VARIETY OF WAYS, TWO OF WHICH ARE ILLUSTRATED. THE CONTACT SEQUENCE, SUCH AS MAKE-BREAK (CONTINUITY) AND BREAK-MAKE (TRANSFER), IS COVERED IN THE APPARATUS FIGURE. DOTS ARE NOT SHOWN AT JUNCTION POINTS OF THE MAKE AND BREAK COMBINATIONS EXCEPT AS ILLUSTRATED IN FIG. 4.

NOTE: FOR WIRE SPRING TYPE RELAYS, SPRINGS ARE NUMBERED IN SETS (POSITION NUMBERS) AND ONLY A SINGLE NUMBER IS SHOWN FOR EITHER A 2 OR 3 SPRING COMBINATION. THE NUMBER IS SO PLACED AS TO IDENTIFY THE FIXED SPRING.

Fig. 3 (right). Basic symbols for functional schematics

designations, (b) functions, (c) features, (d) connecting circuits, (e) alarm routines, (f) how to take equipment out of service, for clearing trouble, etc., and (g) manufacturing test requirements, and 4. reasons for reissue (for revised drawings).

Need for New Drawing

With the ever-increasing complexity of telephone and telegraph switching circuits, especially those for common-control systems, these drawings became progressively more difficult to understand and less suitable for training and maintenance purposes.

It then became necessary, for the more complex circuits, to prepare and maintain a separate series of drawings streamlined principally for use in the training of maintenance personnel. These drawings also were found to be of considerable aid in maintaining the equipment when used as an adjunct to the regular drawings. While this solved the training and maintenance problems, it was found that preparing two different types of drawings for these cases and keeping them in step was costly. To meet this situation the new "all-purpose" drawing described herein was developed.

Analysis of Problems

The more complex circuits are difficult to understand when drawn in the old conventional attached-contact form because of the extreme difficulty of following a circuit path through all apparatus components involved, often requiring tracing through several sheets, and because of the extreme complexity of the network of lines required on the drawing to show these wiring paths.

After careful consideration it was concluded that the proper approach to the problem would be the adoption of a schematic circuit pattern laid out on a straightforward functional path basis, showing each such path in the simplest manner feasible. This required separating individual contacts and coils from relay assemblies and it resulted in the detached-contact method of presentation.

Detaching of contacts is not new. It has been used in various degrees for a number of years both here and abroad. Indeed, in 1917, in the early days of panel machine switching, switching circuits that included power-driven sequence switches were drawn with the cams and contact springs of these switches detached from each other and from the switch structure. No attempt was made at that time to detach relay contacts. The primary purpose of detaching the cams and springs was to place them nearer to the relays with which they were associated, and to avoid serious congestion in the circuit pattern. Each sequence switch in its entirety often had a hundred or more leads connected to it.

Circuit design engineers think in terms of functional circuit paths and for a long time have been preparing their preliminary sketches in detached-contact form. But tradition prevailed and final drawings were always prepared in the conventional attached-contact style until the circuits became so complicated that something had to be done to make them easier to understand.

Objectives

The objectives in the development of the new drawing were as follows:

1. A schematic circuit laid out in patterns of functional circuit paths so that these paths might be readily seen at a glance and easily understood. Complete paths should be shown on the same sheet wherever possible. The contacts and actuating mechanisms of components such as relays should be detached from each other and placed in proper sequence along those functional circuit paths of which they are a part.

2. Provision of separate apparatus figures showing apparatus components in their entirety but without wiring. These figures should specify the apparatus content of an equipment unit. They should also serve as an index for the scattered contacts in the circuit pattern, and show contact sequence of operation.

3. Use of simplified symbols wherever possible.

4. Use of simplified circuit patterns to show well-known functions and features.

5. Use of a sequence chart to show clearly the sequence of operation of relays and other components in the circuit.

6. Use of a uniform drawing sheet of a

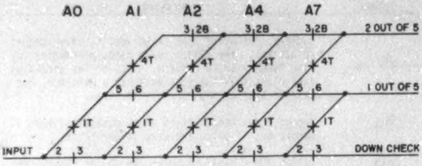
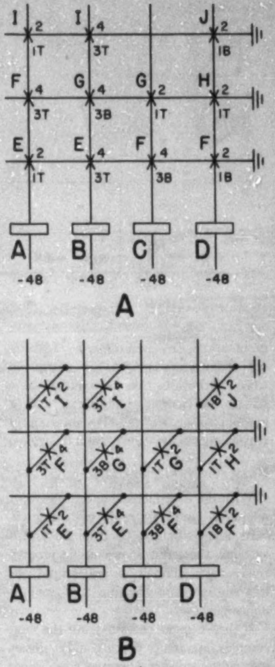


Fig. 4 (left). Symmetric pattern

Fig 5 (right). Relay grid pattern

A—The pattern as used

B—Interpretation of A



size small enough to permit filing in standard binders.

The New Detached-Contact Schematic Circuit Drawing

Basic Differences Between Old and New Types of Drawings

The basic differences in make-up between the two types of drawings may be summed up as follows: The new equivalent of the old attached-contact circuit pattern has been divided into two separate parts: 1. simplified schematic circuit patterns called "functional schematics," and 2. "apparatus figures." Sequence charts have been added and a comprehensive set of indexes used which are the key to the complete drawing. Circuit requirement tables, cabling diagrams, working limits, transmission test requirements, and option indications have been retained. Finally, the format adopted for the new drawings provides for all sheets to be drafted on tracings measuring 22 by 34 inches. When printed for general use all sheets are reduced to 11 by 17 inches. The old drawings were produced in a multiplicity of sizes, mostly large vari-sized sheets, individually folded and unwieldy to use. It therefore follows that the new drawing goes much further than merely detaching contacts.

The accompanying circuit description is similar to that used with the old drawing except that the section on "detailed description of circuit operation" has been simplified considerably. This was made possible by the inclusion of sequence charts.

As in the case of some of the older type of drawings, most of the sheets of detached-contact type of circuit drawings are provided with co-ordinates similar to those used on road maps. These consist of letters placed along two opposite edges of the sheets, and numbers placed along the other two edges. Such co-ordinates are useful in locating specific items of circuit information when they are referred to on other sheets of the drawing, or in the circuit description.

A plan has also been adopted for revising the new drawings on an individual-sheet basis. These drawings are well adapted for this method of handling revisions because changes usually affect fewer sheets than with the older type of drawings. Furthermore, the uniform sheet size plus the inclusion of an index sheet simplifies the work of substituting revised sheets in the complete drawings.

Functional Schematics

Fig. 1 shows a small portion of the schematic circuit pattern of a moderately complex attached-contact drawing for a toll switching system marker circuit which includes approximately 600 relays. It illustrates the complexity of the pattern when drawn in the conventional form. To show the difficulty of identifying and tracing a specific path, the path of the *RL* relay is shown in heavy lines to make it stand out from the large number of intersecting lines that comprise the circuit figure. The *RL* path runs off the sheet seven times and appears on six other sheets of the drawing. The difficulty of tracing such a devious path is obvious.

Simplified Symbols

Fig. 2 depicts the complete *RL* path drawn in the new form, and easily condensed to occupy only a small area of a single sheet. To achieve a maximum degree of simplicity, the old symbols for contacts have been discarded. A make contact (open when not actuated) is represented by an × and break contact (closed when not actuated) by a |. A transfer uses a combination of both. Symbols may be rotated to any position without changing their significance. The two symbols for make and break contacts are the only entirely new ones used. The old battery symbols have been discarded and, instead, the polarity and voltage are shown at the end of each power supply lead. Certain other symbols have been simplified. Fig. 3 includes all of the basic symbols used in the schematic circuit patterns of the new drawing.

Reading a Circuit

Fig. 2 discloses the advantages of the detached-contact technique of circuit presentation when the schematic circuit patterns are laid out on a straightforward functional path basis, emphasizing functional relationships which may be seen at a glance. For example, it is immediately apparent that the operation of the *RL* relay requires that relays *CK* and *AK* be unoperated and that relay *SRL* plus at least one of the 11 relays at the right of the figure be operated. It is also apparent that having operated, the *RL* relay holds through its own locking contact (springs 11 and 12*T*) independent of relays *SRL*, *CK*, and *AK* as long as any one of the other 11 relays remains operated. Functional schematic patterns are made up of circuit paths of this general type, and related paths are grouped together.

Simplified Schematic Circuit Patterns

Several simplified schematic circuit patterns have been adopted. Two of these are geometrical arrangements of contacts which, when once learned, make it possible to identify certain functions immediately. The third is a method of showing repeated patterns in abbreviated form.

The first of these is illustrated by Fig. 4 and is called a "symmetric pattern." In this case the contacts are arranged to fur-

Vol.
Pt. 1
1955

nish one output if two out of five relays are operated at one time. Another output is shown if one relay out of five is operated. A third output, called a "down check," indicates when all relays are down (unoperated). A unique feature of a symmetric pattern is that a "1-story high" pattern indicates a "one-out-of-a-series" check, a "2-story-high" pattern indicates a "two-out-of-a-series" check, etc. The significance of such paths is immediately understood without further analysis. It will be noted that with a pattern of this type, functional designations need be shown but once for each vertical column of contacts.

The second simplified pattern is illustrated by Fig. 5 and is called a "relay grid pattern." This arrangement is intended for use where a group of related paths can be conveniently depicted in a grid pattern. Here each make contact is shown at the intersection of the vertical and horizontal lines and indicates a closure of these lines when the relay is operated. The basic pattern is shown in Fig. 5(A). Fig. 5(B) has been included to show the extra detail which would be required if the use of the contact symbol had been restricted to a single line. The contact symbol shown at the intersection of a vertical and horizontal line is also used to depict closure points on crossbar switches.

The third simplified pattern, called a "repeated circuit pattern," is illustrated by Fig. 6. It shows how to abbreviate repetitions of a pattern without detracting from the clarity of the functional significance.

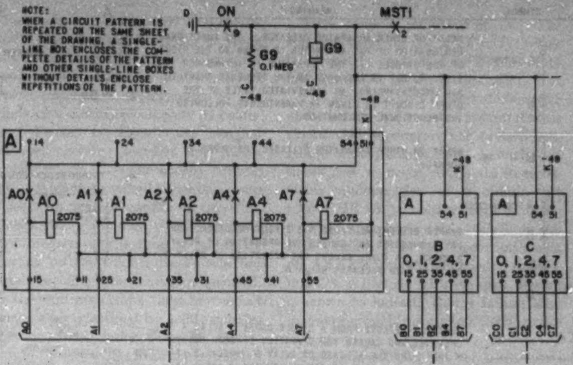
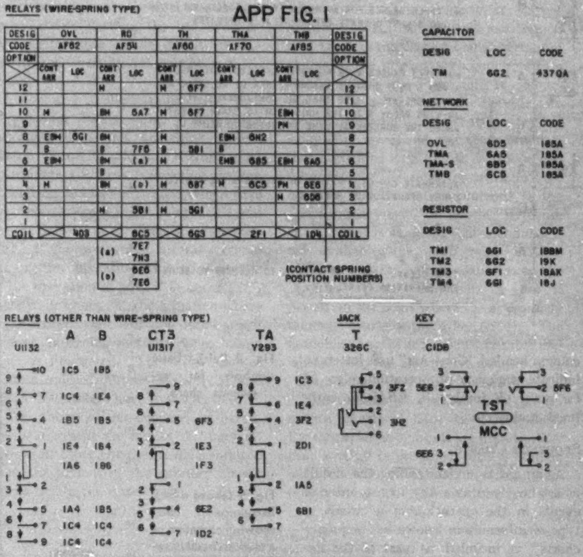

NOTE: WHEN A CIRCUIT PATTERN IS REPEATED ON THE SAME SHEET OF THE DRAWING, A SINGLE-LINE BOX ENCLOSES THE COMPLETE DETAILS OF THE PATTERN AND OTHER SINGLE-LINE BOXES WITHOUT DETAILS ENCLOSE REPETITIONS OF THE PATTERN.

Fig. 6. Repeated circuit pattern

Apparatus Figures

Apparatus figures depict the equipment aspects of a circuit. Just as the completed equipment units are the building blocks of a system, the apparatus figures are the building blocks of the equipment units. This arrangement of apparatus is entirely independent of the functional schematic arrangement; one apparatus figure may include apparatus appearing in more than one functional schematic. Fig. 7 illustrates a typical apparatus figure.


APP FIG. 1

RELAYS (WIRE-SPRING TYPE)

DESIG	OVL		RD		TM		TMA		TMB		DESIG
CODE	AF62		AF54		AF60		AF70		AF85		CODE
OPTION											OPTION
	CONT ARR	LOC	CONT ARR	LOC	CONT ARR	LOC	CONT ARR	LOC	CONT ARR	LOC	
12			M		M	6F7					12
11											11
10	M		BM	6A7	M	6F7			EBM		10
9									PM		9
8	EBM	6G1	BM		M		EBM	6H2			8
7	B		B	7F6	B	5B1	B				7
6	EBM		BM	(a)	M		EMB	6B5	EBM	6A6	6
5			B								5
4	M		BM	(b)	M	6B7	M	6C5	PM	6E6	4
3									M	6D6	3
2			M	5B1	M	5G1					2
1											1
COIL		4D3		6C5		6G3		2F1		1D4	COIL
			(a)	7E7 7H3							
			(b)	6E6 7E6							

(CONTACT SPRING POSITION NUMBERS)

CAPACITOR

DESIG	LOC	CODE
TM	6G2	437QA

NETWORK

DESIG	LOC	CODE
OVL	6D5	185A
TMA	6A5	185A
TMA-S	6B5	185A
TMB	6C5	185A

RESISTOR

DESIG	LOC	CODE
TM1	6G1	18BM
TM2	6G2	19K
TM3	6F1	18AK
TM4	6G1	18J

RELAYS (OTHER THAN WIRE-SPRING TYPE)

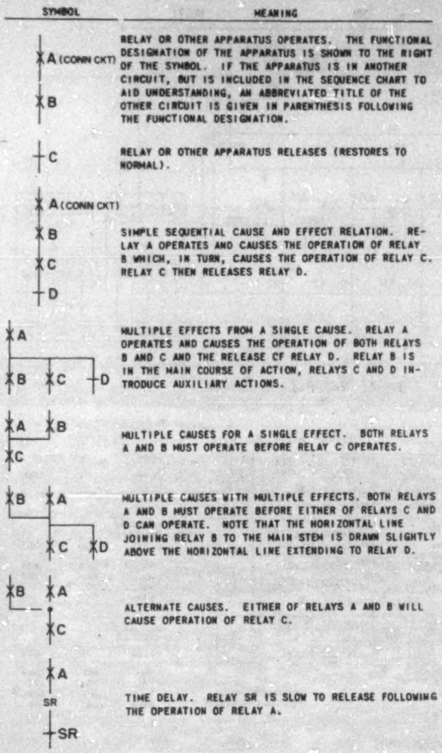
JACK

KEY

Fig. 7. Typical apparatus figure

Each piece of apparatus is identified in the apparatus figure by a code number and also by the same functional designation that it has in the functional schematic. Contact spring and terminal numbers are included. Its co-ordinate location in the functional schematic is also shown. For relays and keys the co-ordinate location is given for each detached part; the apparatus figure serves as an index to the scattered contacts. The absence of a co-ordinate location at a contact indicates that it is a spare or unused contact.

Another function of the apparatus figure is to provide information relative to the type and sequence of operation of each relay contact. Relays of the conventional flat-spring type are symbolized in their entirety, and sequence of operation of contacts is indicated by the symbol configuration. The *U* and *Y* types of relays in Fig. 7 illustrate this method. The new wire-spring type relay contacts are arranged in sets of not over three springs each. Since all sets look alike there is no need for a symbol of an entire relay structure. Instead, information regarding type of contact, sequence of operation, contact position number, etc., is shown in tabular form and is illustrated in Fig. 7 by the *AF*-type relays. The


SYMBOL
MEANING
A (CONN CKT)
B
C
A (CONN CKT)
B
C
D
A
B
C
D
A
B
C
B
A
C
D
B
A
C
A
SR
SR

RELAY OR OTHER APPARATUS OPERATES. THE FUNCTIONAL DESIGNATION OF THE APPARATUS IS SHOWN TO THE RIGHT OF THE SYMBOL. IF THE APPARATUS IS IN ANOTHER CIRCUIT, BUT IS INCLUDED IN THE SEQUENCE CHART TO AID UNDERSTANDING, AN ABBREVIATED TITLE OF THE OTHER CIRCUIT IS GIVEN IN PARENTHESIS FOLLOWING THE FUNCTIONAL DESIGNATION.

RELAY OR OTHER APPARATUS RELEASES (RESTORES TO NORMAL).

SIMPLE SEQUENTIAL CAUSE AND EFFECT RELATION. RELAY A OPERATES AND CAUSES THE OPERATION OF RELAY B WHICH, IN TURN, CAUSES THE OPERATION OF RELAY C. RELAY C THEN RELEASES RELAY D.

MULTIPLE EFFECTS FROM A SINGLE CAUSE. RELAY A OPERATES AND CAUSES THE OPERATION OF BOTH RELAYS B AND C AND THE RELEASE OF RELAY D. RELAY B IS IN THE MAIN COURSE OF ACTION, RELAYS C AND D INTRODUCE AUXILIARY ACTIONS.

MULTIPLE CAUSES FOR A SINGLE EFFECT. BOTH RELAYS A AND B MUST OPERATE BEFORE RELAY C OPERATES.

MULTIPLE CAUSES WITH MULTIPLE EFFECTS. BOTH RELAYS A AND B MUST OPERATE BEFORE EITHER OF RELAYS C AND D CAN OPERATE. NOTE THAT THE HORIZONTAL LINE JOINING RELAY B TO THE MAIN STEM IS DRAWN SLIGHTLY ABOVE THE HORIZONTAL LINE EXTENDING TO RELAY D.

ALTERNATE CAUSES. EITHER OF RELAYS A AND B WILL CAUSE OPERATION OF RELAY C.

TIME DELAY. RELAY SR IS SLOW TO RELEASE FOLLOWING THE OPERATION OF RELAY A.

Fig. 8 (left). Basic symbols for sequence charts

column headed "Cont Arr" uses letters to indicate sequence of operation, e.g., *M* for make, *B* for break, *EBM* for early break make, etc.

Sequence Charts

As an aid to understanding the circuit as a whole and as a key to the order of events in the operation of a circuit, a type of information known as "sequence charts" is included as part of the detached-contact type of circuit drawing. Sequence charts explain circuit action through the use of a compact graphical method, similar in some respects to a family tree or organization chart; in this manner the cause-effect relationship of events that take place in the circuit operation are clearly shown. A sequence chart takes the place of many pages of text material which have been used heretofore to describe circuit operation. Experience has shown that they make it possible to understand the circuit more

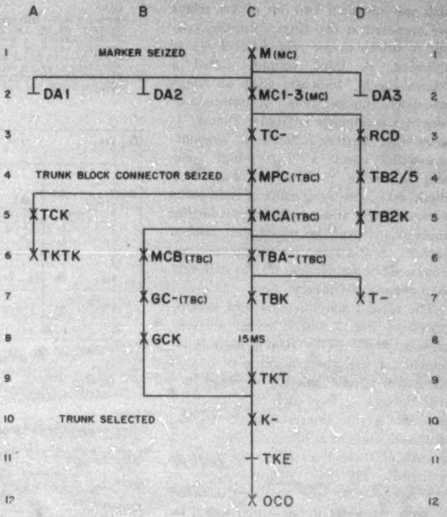

A
B
C
D
1
MARKER SEIZED
M (MC)
2
DA1
DA2
MC1-3 (MC)
DA3
3
TC-
RCD
4
TRUNK BLOCK CONNECTOR SEIZED
MPC (TBC)
TB2/5
5
TCK
MCA (TBC)
TB2K
6
TKTK
MCB (TBC)
TBA- (TBC)
7
GC- (TBC)
TBK
T-
8
GCK
15MS
9
TKT
10
TRUNK SELECTED
K-
11
TKE
12
OCO

Fig. 9 (above right). Sequence chart showing operation of a marker circuit in selecting a trunk

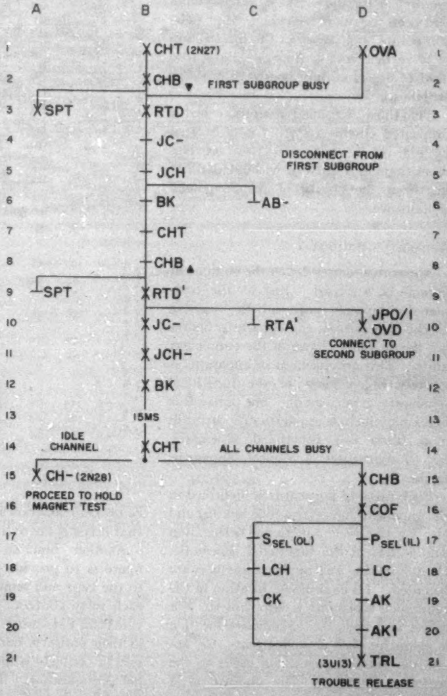

A
B
C
D
1
CHT (2N27)
OVA
2
CHB
FIRST SUBGROUP BUSY
3
SPT
RTD
4
JC-
DISCONNECT FROM FIRST SUBGROUP
5
JCH
6
BK
AB-
7
CHT
8
CHB
9
SPT
RTD'
10
JC-
RTA'
JPO/I
OVD
CONNECT TO SECOND SUBGROUP
11
JCH-
12
BK
13
15MS
14
IDLE CHANNEL
CHT
ALL CHANNELS BUSY
15
CH- (2N28)
CHB
16
PROCEED TO HOLD MAGNET TEST
COF
17
S_{SEL} (OL)
P_{SEL} (IL)
18
LCH
LC
19
CK
AK
20
AK1
21
(3U13)
TRL
TROUBLE RELEASE

Fig. 10 (right). Sequence chart showing the operation of a marker circuit that encounters one or two subgroups in which all channels are busy quickly and easily. In addition, the time and effort required to obtain specific information is greatly reduced, resulting in lower costs for both training and maintenance.

Sequence charts show the direction from cause to effect and are drawn so that the time order of various operations in a circuit is from top to bottom on the page. A characteristic of sequence charts is that the lines indicating the link from cause to effect are read downward or horizontally, but never upward. Fig. 8 illustrates the basic symbols used, and Figs. 9 and 10 show typical sequence charts.

An × indicates the operation of a relay, or magnet, or the activation of an electronic element while the symbol – indicates a release or deactivation. Each event is shown to be the result of a previous event or a combination of events by connecting the appropriate symbols with lines. A single line of action may branch into two or more lines of action and, conversely, several lines of action may converge. The functional designation of each element on the chart is shown beside the symbol representing its action.

In complex switching systems such as those recently developed for automatic local switching, nationwide toll dialing, and automatic message accounting, the individual circuits do not operate independently, but rather actions in one circuit lead to actions and events in associated circuits, and vice versa. The sequence charts therefore cross circuit boundaries in order to present the complete "cause-to-effect" story, and include actions in associated circuits. Where elements of associated circuits are shown, abbreviations enclosed in parentheses are given to represent the name of the external connecting circuit.

It is sometimes necessary to delay particular circuit action to provide sufficient time for other action to be completed. The sequence charts show these delays by skipping one or more time spaces on the vertical line between the action symbols. The actual time delay is shown in this space, or if ordinary slow-operate or slow-release devices are used, the abbreviation *SO* or *SR* is shown.

The action indicated by a sequence chart is based on a set of assumptions making it applicable under the most general conditions. Occasionally, traffic congestion or trouble conditions cause variations in the normal sequence of action. In such cases these variations are described in supplementary sequence charts, or they are shown on the same chart as alternative courses of action.

Since sequence charts show what events should take place and the order in which they should occur, they form a basis for recognizing maloperation in a circuit. A typical example of maloperation is the failure of a circuit to complete its operation. Present-day circuits are designed with self-checking features and major circuits recognize such failures by "timing out" and leaving a record either on a lamp panel or on punched cards. In this way, circuits that encounter trouble indicate their progress at the time of failure by showing on the trouble record whether certain key circuit elements were operated. On the sequence charts, these key circuit elements are identified by solid triangles near the circuit action symbols. Reference to these symbols from the indication on the trouble record shows how far the circuit action had progressed before the trouble was encountered. This greatly assists trouble analysis, since the actions immediately following the last satisfactory performance of the circuit are apparent and their failure can be investigated.

Figs. 9 and 10 illustrate the basic form of sequence charts. Fig. 9 shows the operation of a toll switching system marker circuit in selecting a trunk. At co-ordinate *C*2, lines connecting the operation symbol for the *MC*1-3 relays with the symbols for the *TC-* relay at *C*3 and the *RCD* relay at *D*3 illustrate how one action is shown to cause several succeeding events; operation of *MC*1-3 results in the operation of both *TC-* and *RCD*.

At co-ordinate *C*6, the situation where a particular event will not occur unless several preceding actions have taken place is illustrated. The *TBA-* cannot operate unless both the *MCA* (*C*5) and *TB2K* (*D*5) relays have first operated.

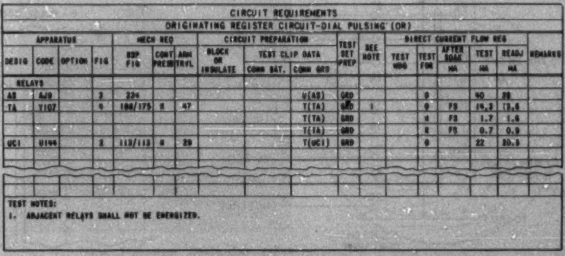
The abbreviations enclosed in parentheses represent the names of "foreign" actions in other circuits as differentiated from the actions of the primary circuit shown on the sequence chart.

Time delay is indicated at *C*8. The specified delay for the *TKT* relay, which is accurately controlled to a time of 15 milliseconds, is shown at this point.

Fig. 10 is a sequence chart showing the operation of a toll crossbar system marker circuit that encounters one or two subgroups of busy transmission channels. It shows how sequence charts aid in recognizing maloperation of a circuit. At relay *CHB* (*B*2) a solid triangle is drawn with its apex downward; this indicates that a failure after this operation would cause a *CHB* trouble indication. At location *B*8, a triangle with its apex upward is shown to indicate that a failure after that action would not cause a *CHB* trouble indication.

Fig. 10 also shows how sequence charts may include alternative assumptions. On lines 1 through 9 for example, the action of the marker when it finds one entire subgroup of channels busy is shown. Since the chart is based on the assumption that there is more than one subgroup of channels, the marker then connects to a second subgroup. From line 14 the action of the marker takes either of two courses, depending on what further assumptions are made. If it is assumed that a channel is available on the second subgroup, the *CH-* relay (co-ordinate *A*15) operates and the action proceeds to the "hold magnet test" on another sequence chart. If it is assumed that a channel is not available, the *CHB* relay at *D*15 operates and the action continues to the "trouble release" function. A dot is placed on the point where the action may take one course or the other (*B*14). The branching lines of action are separated and labeled with the alternative assumptions, "idle channel" and "all channels busy."

The number of symbols used on sequence charts is kept to a minimum. For example, no symbols are used to show that a relay is provided with a holding path following its operation, since no action is involved. Circuit details of this



CIRCUIT REQUIREMENTS

ORIGINATING REGISTER CIRCUIT-DIAL PULSING (OR)

APPARATUS				MECH REQ			CIRCUIT PREPARATION			TEST SET PREP	SEE NOTE	DIRECT CURRENT FLOW REQ					REMARKS
DESIG	CODE	OPTION	FIG	BSP FIG	CONT PRESS	ARM TRVL	BLOCK OR INSULATE	TEST CLIP DATA				TEST HDG	TEST FOR	AFTER SOAK	TEST	READJ	
								COMM BAT.	COMM GRD					MA	MA	MA	
RELAYS																	
AS	AJ9		3	234					U(AS)	GRD			0		40	28	
TA	Y107		6	166/175	8	47			T(TA)	GRD	1		0	FS	14.3	13.6	
									T(TA)	GRD			N	FS	1.7	1.6	
									T(TA)	GRD			R	FS	0.7	0.9	
UCI	U144		2	113/113	8	39			T(UCI)	GRD			0		22	20.5	

TEST NOTES:

1. ADJACENT RELAYS SHALL NOT BE ENERGIZED.

Fig. 11. Arrangement of information in circuit requirement table

type are shown on the functional schematics. As a result, sequence charts are free to depict the progression of circuit action in concise symbolic notation. The simple form of these charts is an important factor contributing to their marked success in both training and use on the job.

Circuit Requirement Tables

Circuit requirement tables give the electrical and mechanical requirements of such items as relays, registers, and keys. This information is required when testing or readjusting the apparatus, to ensure the proper operation of the circuit. Information is also included for preventing interference with connecting circuits, operation of alarms, and blowing of fuses when making tests or adjustments. Apparatus is identified in the tables by both its functional designation and its code number. A typical circuit requirement table is shows by Fig. 11.

Cabling Diagrams

The cable connections between a circuit and other parts of the plant are shown in cabling diagrams. These are divided into figures as required. Direct reference is established in the cabling diagram to the related apparatus figure. Fig. 12 shows a typical cabling diagram.

Notes

There are five different series of notes used on the new drawing. Four of these are usually shown on a single sheet (or group of sheets) of the drawing. These are: 1. circuit notes, 2. equipment notes, 3. information notes, and 4. cross-connection information notes. A fifth series, known as "sheet notes," may be used on any sheet of the drawing and their application is limited to the sheet on which they appear.

Circuit notes cover such items as fusing, summary of features and options with a list of apparatus and wiring required for each, job record information, and other items pertaining to the circuit.

Equipment notes are intended primarily for use in preparing manufacturing information. They cover such items as size, resistance, and length of certain critical leads, and mounting restrictions of apparatus.

Information notes cover information of a general nature which does not fall under the categories of the other series of notes. They usually cover information describing a function of an external circuit where this affects the operation of the circuit on which the note is shown. Typical examples are rated time intervals of an interrupter shown on another drawing, and reference to arrows indicating direction of transmission as in 4-wire repeaters.

Cross-connection information is required primarily for making cross connections that are subject to periodic changes which are made by the maintenance personnel. The information is usually shown in tabular form. Explanatory notes are used when required and these notes sometimes include schematic figures for clarification.

Indexes

A comprehensive set of indexes is used and is the key to the entire drawing. These consist of: 1. sheet index, 2. apparatus index, 3. lead index, 4. option index, and 5. sequence chart index. The first four of these indexes appear on the first sheet or sheets of the drawing. Also included on the first sheet is a table of supporting information such as associated equipment drawings and specifications. The sequence chart index is shown in that part of the drawing covering sequence charts. Indexes are illustrated in Fig. 13. A sheet index shows the contents of each sheet of the drawing.

The apparatus index lists all of the component apparatus used in the circuit. Information is divided into types of apparatus such as relays, capacitors, resistors, and inductors. Apparatus items are identified by functional designations followed by the sheet number and coordinate location of their appearance in the functional schematic. For items which have parts detached, such as relays, keys, etc., the location in the functional schematic refers only to the coils or actuating mechanisms. The location of detached parts is shown only in the apparatus figures. Also included is the apparatus figure number of which the item

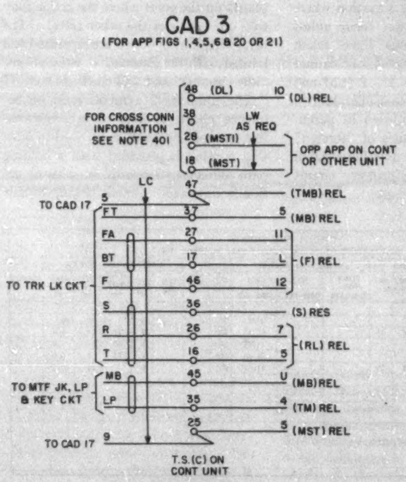
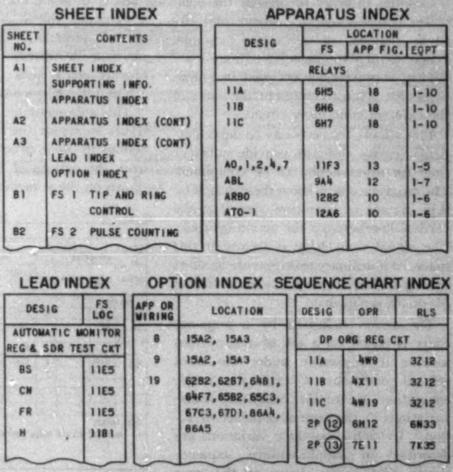


Fig. 12. Typical cabling diagram



SHEET INDEX

SHEET NO.	CONTENTS
A1	SHEET INDEX SUPPORTING INFO. APPARATUS INDEX
A2	APPARATUS INDEX (CONT)
A3	APPARATUS INDEX (CONT) LEAD INDEX OPTION INDEX
B1	FS 1 TIP AND RING CONTROL
B2	FS 2 PULSE COUNTING

APPARATUS INDEX

DESIG	LOCATION		
	FS	APP FIG.	EQPT
RELAYS			
11A	6H5	18	1-10
11B	6H6	18	1-10
11C	6H7	18	1-10
A0,1,2,4,7	11F3	13	1-5
ABL	9A4	12	1-7
ARB0	12B2	10	1-6
AT0-1	12A6	10	1-6

LEAD INDEX

DESIG	FS LOC
AUTOMATIC MONITOR REG & SDR TEST CKT	
BS	11E5
CN	11E5
FR	11E5
H	11B1

OPTION INDEX

APP OR WIRING	LOCATION
8	15A2, 15A3
9	15A2, 15A3
19	62B2,62B7,64B1, 64F7,65B2,65C3, 67C3,67D1,86A4, 86A5

SEQUENCE CHART INDEX

DESIG	OPR	RLS
DP ORG REG CKT		
11A	4W9	3Z12
11B	4X11	3Z12
11C	4W19	3Z12
2P ⑫	6H12	6N33
2P ⑬	7E11	7K35

Fig. 13. Example of indexes

is a part, and the location where it is mounted on the frame.

The lead index shows the designations of all leads leaving a circuit together with the sheet number and co-ordinate location on the sheet where they appear. The name of the connecting circuit is also shown.

The option index gives the location of all optional apparatus or wiring, or both, in the functional schematic. If apparatus is involved, the apparatus figure number is given.

The sequence chart index summarizes, for quick reference, the sheet number and co-ordinate location on the sequence chart of the various "operate" and "release" actions of relays, switches, etc.

Conclusions

Experience, to date, with the new drawing has been most satisfactory. Substantial savings in both design and drafting effort are being realized. While it is a little early accurately to evaluate its effectiveness in all of the various fields of use, present indications are most promising. In the meantime all of the circuits for new systems and major revisions of the more complex circuits of existing systems are being prepared in the new form.

Thus, a troublesome situation has been met with a truly "all-purpose" drawing. But the job will never end. With the thousands of drawings in use throughout the Bell System and the hundreds of new and revised drawings being prepared each year, carefully planned standardization control must be continuously maintained. In this way the many problems which arise from time to time are solved either by modifying existing practices or by adopting new ones.

The Design of a New 2-Element Single-Disk Polyphase Meter

W. J. SCHMIDT
NONMEMBER AIEE

Synopsis: A new 2-element single-disk polyphase meter has been designed which incorporates a novel method of compensation to correct for interelement interference and a new approach to the problem of providing means for calibrating the meter. The method used to provide compensation for interelement interference and the design of the adjusters for calibrating light load and phase balance are described.

THE art of metering polyphase systems started with the Shallenberger design of 1896 and the Thomson design of 1899.[1] It is interesting to note that both of these designs contained two driving elements acting on a single disk. Because of the interference, however, between the fields of the two elements acting on a common disk, this construction was abandoned around the turn of the century in favor of multidisk meters. A Davis and Conrad[1] patent in 1898 disclosed the first multidisk polyphase induction meter to operate correctly on unbalanced loads. As the polyphase metering art progressed, manufacturers turned their efforts toward producing a more compact, simpler, and less expensive polyphase meter. These efforts have followed two channels; development of the multidisk meter and development of the single-disk meter incorporating means for eliminating the interference.

The new polyphase meter described in this paper, Fig. 1, was designed to provide a long-life meter with extended overload range and built-in protection against lightning surges. These features which are now available in single-phase watthour meters were combined with compactness and ease of calibration. Compactness of design was achieved by making use of the single-disk design and a unique method of compensating for interference proposed by B. E. Lenehan, Westinghouse Electric Corporation, Newark, N. J. Ease of calibration was achieved by providing independent adjustments designed specifically for use on a 2-element polyphase meter.

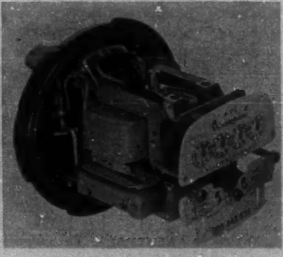


Fig. 1. New 2-element single-disk polyphase meter

The design of the new 2-element single-disk polyphase meter will be described. Since the means for providing a long-life meter with extended overload range and built-in protection are generally known to the art, only three main phases of the design will be discussed:

1. The way compensation is provided to eliminate interference and to provide a compact design.
2. A simplified means for adjustment.
3. A phase-balance adjuster that does not change the relationship between full load, light load, and power factor.

Interference in a 2-Element Meter

A short discussion may be in order to show how interference arises in a single-disk 2-element polyphase meter. To simplify the explanation, interference may be thought of as arising from three separate interactions between eddy currents in the disk and fluxes from the electromagnet poles. These interactions are cross effects and occur simultaneously between

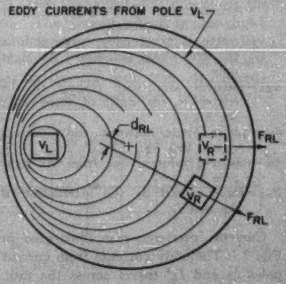


Fig. 2. Voltage Interference

Paper 55-483, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 25, 1955; made available for printing June 1, 1955.

W. J. SCHMIDT is with the Westinghouse Electric Corporation, Newark, N. J.

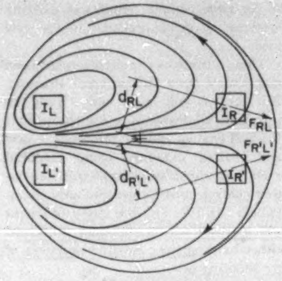


Fig. 3. Current interference

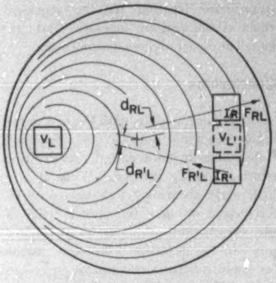


Fig. 4. Voltage-current interference

the poles of the two elements when they are energized. The interactions are as follows:

1. Voltage interference, voltage pole flux interacting with eddy currents from the opposite voltage pole.
2. Current interference, current pole flux interacting with eddy currents from the opposite current pole.
3. Voltage-current interference, current pole flux interacting with eddy currents from the opposite voltage pole.

Voltage inteference is illustrated in Fig. 2. The eddy currents from voltage pole V_L travel through the disk in circular paths and interact with the flux from voltage pole V_R. This interaction produces a torque according to the equation

$$T_{V_{LR}} = T_{V_L V_R} = k i_L \Phi_R d_{RL} \cos \beta_{RL} \tag{1}$$

where

$T_{V_{LR}}$ = total voltage interference torque caused by the interaction between eddy currents from the left voltage pole and flux from the right voltage pole
$T_{V_L V_R}$ = voltage interference torque caused by interaction of eddy currents from the left voltage pole interacting with the flux from the right voltage pole
k = constant dependent on the physical arrangement of components (best arrived at empirically)
i_L = effective value of eddy currents from the left voltage pole under the right pole
α_{RL} = moment arm, with respect to the center of the disk, of force F_{RL}; see Fig. 2
β_{RL} = phase angle between eddy current i_L and flux Φ_R
Φ_R = effective value of flux from the right pole
F_{RL} = force caused by the interaction of Φ_R and i_L

Current interference is illustrated in Fig. 3. The eddy currents from current poles I_L and I_L' travel across the disk, as shown in Fig. 3, and interact with the flux from current poles I_R and I_R'. This interaction produces a torque according to equation 1 in the following manner

$$T_{I_{LR}} = T_{I_L I_R} - T_{I_{L'} I_{R'}} \tag{2}$$

Voltage-current interference is illustrated in Fig. 4. The eddy currents from voltage pole V_L travel across the disk and interact with the fluxes from current poles I_R and I_R'. This interaction produces a torque according to equation 1 in the following manner

$$T_{VI_{LR}} = T_{V_L I_R} + T_{V_L I_{R'}} \tag{3}$$

The torques produced by all of these interactions are dependent on the phase angle between the eddy currents from one element and the flux from the other element (cos β). If the meter is to read true power, the torque on the disk must be produced only by interactions within each element according to the following

$$T_{VI_R} = k i_R \Phi_R d_R \cos \beta_R \tag{4}$$

The total torque on the meter, however, is equal to

$$T_{VI_R} + T_{VI_L} + T_{V_{LR}} + T_{I_{LR}} + T_{VI_{LR}} + T_{V_{RL}} + T_{I_{RL}} + T_{VI_{RL}} \tag{5}$$

For the meter to register true power, all of the interfering torques in equation 5 will have to be zero.

Interference Correction in a 2-Element Meter

The voltage and current interference described in the preceding section is corrected by arranging the watt-hour meter elements symmetrically with respect to the center of the disk. Voltage interference is corrected by a symmetrical arrangement which causes force F_{RL} (see dotted pole in Fig. 2) to fall in line with the center of the disk reducing d_{RL} to zero in equation 1. Current interference is also corrected by this arrangement, as can be seen by reference to Fig. 3. Since corresponding poles (I_R and I_R' or I_L and I_L') are parts of the same magnetic circuits, their strengths and the strengths of their eddy-current fields are equal. Symmetrical arrangement also makes d_{RL} equal to $d_{B'L'}$, thus $I_{I_{L'} I_{R'}}$ in equation 2 and the total torque $I_{I_{LR}}$ are zero.

Another interference problem is that of voltage-current interaction. This is eliminated by a compensating winding applied to the voltage poles. Fig. 5 illustrates how these windings are connected on a 1-, 2-, or 3-phase 3-wire meter.

To illustrate how this winding operates, assume that the effect of the eddy currents in the disk is the same as superimposing a new pole V_L' (shown dotted in Fig. 4) on the voltage pole V_R, with a strength equal to CV_L. C is a constant equal to approximately 0.02 and is determined by the amount the disk attenuates the eddy currents from pole V_L in the area

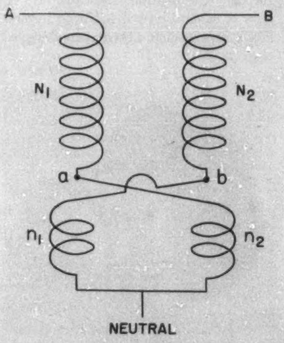
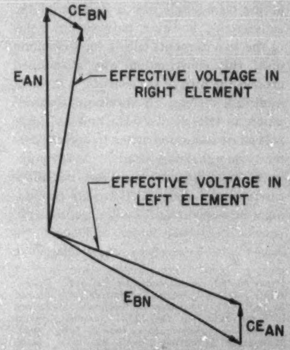


Fig. 5. Potential coil windings for voltage-current interference compensation



CE_{BN}
E_{AN}
EFFECTIVE VOLTAGE IN RIGHT ELEMENT
EFFECTIVE VOLTAGE IN LEFT ELEMENT
E_{BN}
CE_{AN}

Fig. 6. Voltage relationship in an uncompensated meter

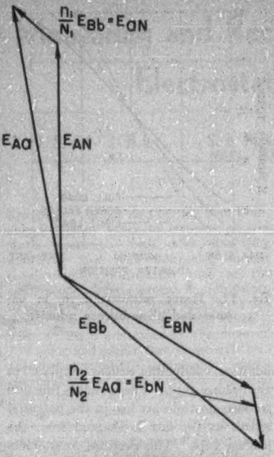

$\frac{n_1}{N_1}E_{Bb} = E_{aN}$
E_{Aa}
E_{AN}
E_{Bb}
E_{BN}
$\frac{n_2}{N_2}E_{Aa} = E_{bN}$

Fig. 7. Effect of compensating winding on voltage relationship

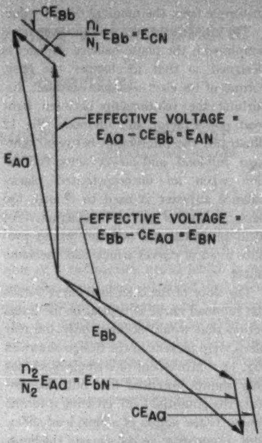

CE_{Bb}
$\frac{n_1}{N_1}E_{Bb} = E_{CN}$
EFFECTIVE VOLTAGE = $E_{Aa} - CE_{Bb} = E_{AN}$
E_{Aa}
EFFECTIVE VOLTAGE = $E_{Bb} - CE_{Aa} = E_{BN}$
E_{Bb}
$\frac{n_2}{N_2}E_{Aa} = E_{bN}$
CE_{Aa}

Fig. 8. Voltage relationship in a compensated meter

occupied by the pole V_R. This has the same effect as adding the phasor CE_{BN} to the phasor E_{AN} creating a new effective voltage in the right element equal to $E_{AN}+CE_{BN}$. The new value of effective voltage $E_{AN}+CE_{BN}$ creates a torque in the right element equal to $I_{VI_R}-I_{VI_{LR}}$ (equation 4−equation 3). Fig. 6 shows the voltage relationship in an uncompensated meter.

The compensating windings function to eliminate voltage-current interference by producing a voltage $n_1/N_1E_{BN}=E_{aN}$, Fig. 7, which is equal and opposite to CE_{BN}. The final effective voltage, therefore, is $E_{Aa}-CE_{BN}=E_{AN}$ nullifying the voltage-current interference, as shown in Fig. 8. Voltage E_{aN} will always be equal to and 180 degrees out of phase with CE_{Bb} so that the compensation will work for any phase relationship between E_{AN} and E_{BN}.

The Design of Simplified Adjustments

Conventional polyphase meters have always provided a light-load adjuster and a phase-balance adjuster on each electromagnet. This practice is probably a carry-over from the early designs which were essentially two or more single-phase electromagnets arranged to drive a single disk or multiple disks on a common shaft. The light-load and phase-balance adjusters discussed herein are designed to perform the specific functions of calibrating light load and phase balance in a 2-element polyphase meter. Designing for a specific job allowed the use of only one light-load and one phase-balance adjuster; this led to a simplification in the adjustment of the meter. The details of the design of the two adjusters will be discussed.

Light-Load Adjuster Design

In an induction watt-hour meter, the light-load slider produces a fixed compensating torque to provide correction for three factors which influence the accuracy of the meter at the arbitrarily chosen light-load calibration point, generally 10 per cent of the meter name-plate rating. The following will cover each of these factors individually and show why only one light-load slider is necessary on a 2-element meter.

The first factor which requires correction is "creep," which is a constant torque effect created by dissymmetries in the potential circuit of the electromagnet. These dissymmetries may be electrical or mechanical in nature and produce a constant torque tending to turn the disk. Since all of the potential coils of a polyphase meter are energized in service, the total creep effect produces a constant torque on the disk. Only one light-load slider is necessary to produce an equal and opposite torque.

Frictional effects in the baring system, although extremely small, produce enough torque at light loads to require compensation. Friction torque has an equal influence on the basic torques of all electromagnets since they all act on the same moving element. As a result, friction may be compensated for by adding a single constant torque to the moving element.

The third factor which influences the accuracy of the meter at light loads is the difference in the inherent responses of any two electromagnets on varying loads. Fig. 9 illustrates this effect. Although both electromagnets are arranged to produce the same torque at 100 per cent of rated load, their torques are not identical at 10-per-cent load. In the example given, electromagnet *A* has a registration (registration is proportional to torque) of 99 per cent, and electromagnet *B* has a registration of 98.5 per cent at 10-per-cent load. Fig. 10 illustrates the change in the two curves when a light-load slider on electromagnet *A* is adjusted so that the registration of electromagnet *A* is 100 per cent at 10-per-cent load. As shown by the curve, the registration of electromagnet *B* is still 0.5 per cent low. It would appear that the adjustment of a light load slider on electromagnet *B* would increase its registration by 0.5 per cent so that it corresponded to the registration of electromagnet *A*. This is not the case. The torque added by the light-load slider on electromagnet *B* would also be added to electromagnet *A* and it would then be 0.5 per cent fast. No matter how many light-load sliders there are, the

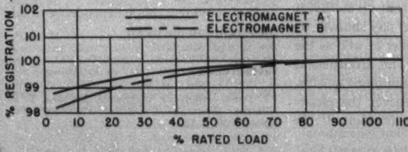


Fig. 9. Inherent response of electromagnets to varying loads

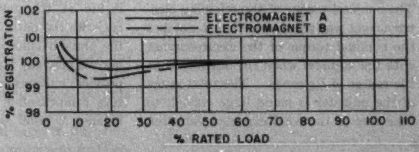


Fig. 10. Inherent response with light-load slider compensation added

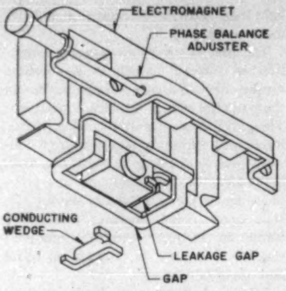


Fig. 11. Mechanical arrangement of phase-balancer parts

0.5-per-cent difference between the electromagnets cannot be eliminated. One light-load slider on electromagnet A will compensate for this effect on both electromagnets provided it has sufficient range of adjustment. In actual practice the light-load point would be calibrated with both elements energized and electromagnet A would be 0.25 per cent fast and electromagnet B would be 0.25 per cent slow.

Phase-Balance Adjuster Design

The phase-balance adjuster is provided on multielement meters to equalize the basic torques of all the electromagnets at some common load point, usually 100 per cent of rated load. The phase-balance adjuster provides a means for varying the basic torque of the electromagnet on which it is mounted.

Balancing the torques of the electromagnets in a 2-element meter is simplified if the torque of one of the electromagnets can be used as a reference and the torque of the other electromagnet adjusted to match the reference torque. The foregoing statement outlines the basic thinking which guided the design of the phase-balance adjuster. Phase-balance adjustment is achieved by shunting potential flux away from the disk gap by means of a mechanism, as shown in Fig. 11. This adjuster is mounted on the adjustable electromagnet and a fixed shunt is mounted on the reference electromagnet. The nominal torque of the reference electromagnet was reduced so that it matches the nominal torque of the electromagnet with the adjuster when the adjuster is at the center of its range. The total range of the adjuster is made large enough to take care of the total expected variation in torque from the nominal design value.

The adjustment of a polyphase meter is simplified if the phase-balance adjuster is designed so that it changes the basic torque of its electromagnet without disturbing the relationship between light load, power factor, and full load. Fig. 12 illustrates the interaction between light-load, full-load, and power-factor calibration when an uncompensated phase-balance adjuster is used to change the basic torque of an electromagnet. An ideal phase-balance adjuster would produce a set of curves which had the same slope.

The slope of the light-load curve versus the full-load curve is typical of the interaction that would occur if only the rear block were present in the design shown in Fig. 11. Movement of a single block not only shunts potential flux but also distorts the potential field, creating a torque similar to the action of a light-load slider. This torque changes the slope of the light-load curve relative to the full-load curve. This condition is corrected by adding a second block symmetrically located with respect to the first block. The two blocks move with a turn-buckle action when the adjuster screw is turned. As a result the blocks are always symmetrically located with respect to the center line of the electromagnet, thus maintaining symmetry of the potential flux field.

The slope of the power factor curve versus the full-load curve, as shown in Fig. 12, illustrates the interaction that would result with uncompensated phase-balance adjusters. As the phase balancer is moved to shunt more potential flux, power factor calibration slows down relative to full-load calibration. This change is a result of the following effects: the quadrature loop and disk are cut by less potential flux, thus decreasing the quadrature compensation of these components; part of the potential flux, passing through the leakage gaps, is shunted from a low watt-loss path in the electromagnet to a higher loss path in the phase-balance adjuster, thus increasing the watt loss in the potential circuit and decreasing power-factor registration.

To correct the conditions just described, two forms of compensation are added to the phase balancer. As shown in Fig. 11, the quad is bent up so that it encircles the path of the flux shunted by the phase-balance adjuster as well as the normal disk gap flux. This arrangement provides a means for maintaining quadrature compensation whenever the phase-balance adjuster setting is changed. In addition, conducting wedges are placed in the leakage gaps. These wedges function to decrease the watt loss in the potential circuit as the flux is shunted from the leakage gaps. This decrease counteracts the increase caused by the shunted flux traveling in the phase-balance adjuster. By combining these two compensations, the slope of the power-factor curve is made to match the slope of the full-load curve in Fig. 12.

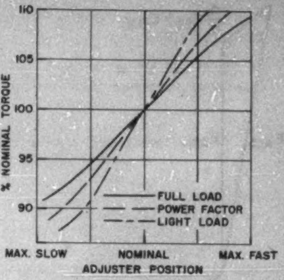


Fig. 12. Torque relationships in an uncompensated phase-balance adjuster

Conclusion

The art of metering polyphase circuits has developed to the point where it is possible to design a compact 2-element polyphase meter that is not only highly accurate but also simple to calibrate. Methods of compensating for interelement interference are described. A compact 2-element polyphase meter is the result of the application of these methods of compensation.

Simplified adjustment of a 2-element polyphase meter is achieved by using one light-load adjuster and one phase-balance adjuster. This simplification is possible when the factors involved in the design of the adjusters are evaluated by the methods described in this paper. Adjustment of the meter is further simplified by a novel design of phase-balance adjuster which changes the basic torque of the electromagnet without disturbing the relationship between light-load, power-factor, and full-load calibrations. Methods of providing this independence of adjustment are discussed.

Reference

1. DEVELOPMENT OF WATTHOUR METERS, T. D. Barnes. *Westinghouse Engineer*, Pittsburgh, Pa., Nov. 1950, p. 230.

Measuring and Recording Atmospheric Electrostatic Potential

J. S. CARROLL
FELLOW AIEE

S. B. HAMMOND
ASSOCIATE MEMBER AIEE

E. H. STEWART
ASSOCIATE MEMBER AIEE

Synopsis: Various means have been available for many years to measure electrostatic potentials with ever increasing sensitivity and with decreasing current requirements. Instruments such as the electroscope or the common laboratory electrostatic voltmeter have been developed to the point where they operate reliably, with an input resistance of the order of 10^{16} ohms and higher, but in general these extremely sensitive instruments cannot conveniently be used to record potential over a period of time, thus requiring the continuous presence of an operator. A type of electrostatic voltmeter is described which is designed to record potentials as low as 1 volt d-c with an input resistance of 10^{14} ohms using an electrostatic generator to provide an a-c signal proportional to voltage, and a simple audio amplifier requiring only a commercial 1-milliampere recording meter in the way of auxiliary equipment.

MEASUREMENT of atmospheric electrostatic potentials has intrigued scientists and engineers for many years. The principal difficulty encountered in measuring these potentials has been similar to the essential problem of many other types of electric instrumentation, namely, the power required to operate the instrument distorts the potential thereby measured. The instrument reading, therefore, is not the desired potential but something different which must be corrected to obtain the true value. This problem is very difficult when measuring atmospheric electrostatic potential because of the extremely small amount of power available to drive the measuring instrument. This power limitation requires extremely high input resistance. The development of electroscopes and other instruments with input resistance of 10^{12} ohms and higher has been necessary to obtain reasonably reliable measurements. Carroll and Hammond[1] present a discussion of this problem while describing an instrument suitable for taking these readings but not suitable for continuous recording without an operator.

The instrument described herein consists of a radioactive probe which measures the atmospheric electrostatic potential, a rotating capacitor which converts the d-c signal from the probe to an a-c signal, an amplifier to amplify this signal, and a recording meter to record it. These basic parts are shown in Fig. 1. The important requirements specified were:

1. The instrument must be reliable, portable, and maintain calibration, using commercially available components as much as possible.
2. The leakage current must be low.
3. Recording and indicating must be done on standard 1-milliampere d-c instruments.
4. Equipment must be capable of operation by people not especially trained in electrical measurement technique.

The Probe

Measurement of electrostatic potential is usually accomplished by using a well-insulated probe which can assume the potential of the surrounding atmosphere. The probe in turn is connected to a voltage-measuring device. A simple needle-point probe could be used if it were well enough insulated, but its response would be slow. To give quicker response, and with less meticulous insulation, the historic flame probe or water-drop probe could be used, but they are relatively cumbersome. To obtain rapid response while retaining a convenient configuration, a radioactive probe was initially used for this instrument. This is a needle probe, Fig. 2, which is made radioactive by wrapping a piece of polonium foil around most of the needle. The preparation of this foil was similar to the preparation described by Martin.[2] The polonium is an alpha emitter and provides a region of highly ionized air around the probe with a maximum diameter of approximately 8 centimeters. The presence of this conducting region permits rapid equalization of any potential difference that may exist between the probe and the surrounding atmosphere, and permits rapid response to changes in potential. The conducting region should not distort the electrostatic field, provided the field is relatively large, because of the balance of charges of the ions generated. Further work is being done on the effects of probe configuration and variations of raidoactivity.

There are theoretically several possible causes of a discrepancy between probe voltage and the voltage in the surrounding medium. Probably the best known of these arises from the drift velocity of ions (as differentiated from their random kinetic velocity), and this has caused considerable difficulty in the use of probes in partially evacuated tubes. In the case of these gas tubes, appreciable ion velocities can be attained and this ion stream will distort the probe reading because of impingement of the ions on the probe. This situation does not occur when measuring in air, since the drift velocity of the ions is negligible compared to their random velocity, even when subjected to strong electrostatic fields. Other possible causes of error such as the probability of more ions of the predominant polarity impinging by random kinetic velocity on the probe and the effect of the work function of the surface of the probe are also negligible when using the probe in the ambient atmosphere and measuring potentials of 1 volt or greater.

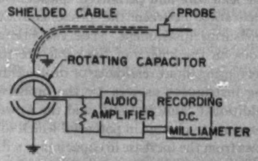


Fig. 1. Functional diagram of rotary electrostatic voltmeter

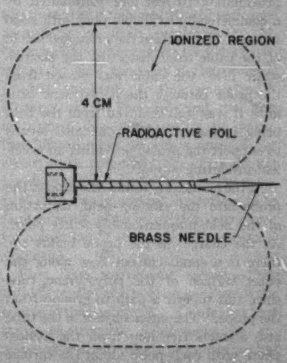


Fig. 2. Radioactive needle probe

Paper 55-489, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 30, 1955; made available for printing May 5, 1955.

J. S. CARROLL is with the Stanford University, Stanford, Calif.; S. B. HAMMOND is with the University of Utah, Salt Lake City, Utah, and E. H. STEWART is with the Wesix Electric Heater Company, San Francisco, Calif.

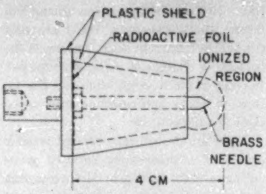


Fig. 3. Radioactive needle probe with shield

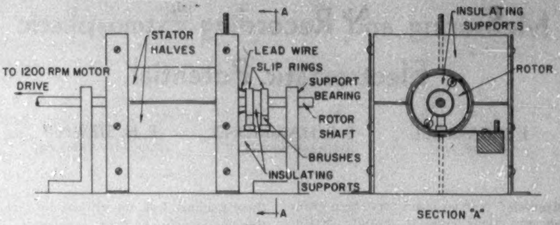


Fig. 4. Rotating capacitor

A probe used to measure atmospheric potential must be employed to give minimum distortion of the electrostatic field which exists in the absence of the probe. It is necessary to use a shielded conductor from the probe to the measuring instrument since it is desirable to separate the measuring instrument from the probe by several feet to minimize the field distortion by the instrument case. To accomplish this, a length of coaxial cable was slipped inside a 6-foot length of polystyrene tube. The shield was terminated and left unconnected inside the tube about 1 inch from the probe and was grounded to the instrument at the terminal end. The shielding provided by the cable effectively prevented pickup of random electrostatic or electromagnetic fields by the lead wire and permitted considerable separation between the probe and the measuring instrument. The two limitations on the length of the cable are the decreasing input resistance of the measuring instrument as seen by the probe as additional cable is added and the slower response to change in potential which follows from the increase in capacitance with increase in cable length. The authors used a commercially available cable approximately 10 feet long terminated by a commercially available plug with good success. To restrain the grounded shield of the cable from distorting the electrostatic field, the cable was passed from the probe through the polystyrene tube until it was well removed from the field being measured. The tube also served to support the probe, providing an excellent insulating support.

At the moment of initial insertion of the probe into the existing field, distortion of the field unquestionably takes place. As the charge on the probe builds up, there is a small current flow along the outer surface of the polystyrene tube since this tube is a path to ground from the probe. The outer surface of the tube also attracts free ions from the atmosphere until each point along the surface of the tube is at the potential which that point in space had achieved before the probe was inserted into the field. When this equilibrium of charge is achieved, the electrostatic field is no longer distorted. In the ideal case where no current is drawn, this equalization would be perfect. In this instrument the impedance to ground measured at the probe is approximately 10^{14} ohms and for the measurements taken by the authors it was considered that no appreciable distortion of the field took place.

One other point to be taken into account in considering field distortion is the size of the probe in relation to the voltage gradient and the field being measured. The probe shown in Fig. 2 is small when compared to most outdoor fields, and can be used successfully for many indoor measurements. This probe must be considered as a conducting cylinder, 3 centimeters long and 8 centimeters in diameter. The large diameter is the ionized zone, which is sufficiently conducting to be an equipotential volume, and its radius is the radius of the probe plus the range in air of the alpha radiation from the foil (about 3.8 centimeters).

In a small enclosure, a probe as large as this distorts the field too much, and a modified design is shown in Fig. 3. The plastic shield in this design stops all radiation except that aimed at a small volume close to the point of the probe. The conducting volume is a cylinder approximately 1/2 inch in diameter and 1/2 inch long, making it possible to use this probe in much smaller rooms than was possible with the probe of Fig. 2. The plastic shield serves to minimize field distortion by accumulating a charge in the same way the polystyrene tubing shields the coaxial cable.

Rotating Capacitor

The rotating capacitor is basically a type of electrostatic generator, connected to deliver alternating current instead of the more usual direct current. Kirkpatrick[3-6] was successful in developing this principle to measure the high target voltages used on X-ray tubes when conventional instruments of that time required more power than was available.

Physically, this portion of the instrument, Fig. 4, consists of two concentric brass cylinders approximately 2 inches in diameter and 4 inches long, with about 1/16 inch separating the inner and outer cylinders. Both cylinders are split lengthwise, and the two halves are well insulated from each other. The inner cylinder halves are mounted on a shaft, but insulated from it, and are free to rotate inside the fixed, outer cylinder halves. The inner half-cylinders are connected to slip rings which in turn make contact with brushes connected to the amplifier. The rotor part is driven by a synchronous-type motor operating at 1,200 rpm.

In operation, the probe is connected to one fixed plate while the other plate of the stator is grounded, thus establishing an electrostatic field between the two fixed plates. A charge is separated between the two rotating plates by this field. This charge alternates between the rotor halves as the rotor revolves, flowing through the resistor conected between the two brushes, thus creating an alternating voltage.

The theoretical magnitude of the output signal is expressed as

$$I = \frac{\pi}{\sqrt{2}} CVn$$

where

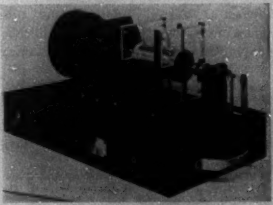


Fig. 5 Rotating capacitor assembly

Vol.
Pt. 1
1955

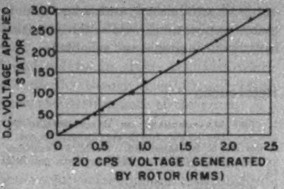


Fig. 6. Input voltage–output voltage relationship of rotary capacitor

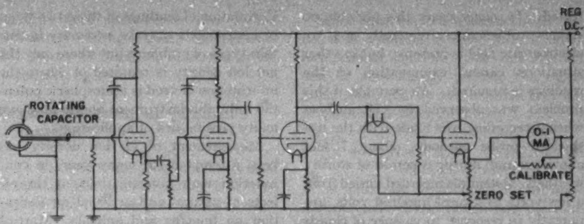


Fig. 7. Amplifier schematic diagram

I = the rms value of current
C = the capacitance from one plate of the stator to the half of the rotor facing it
V = the voltage applied to the plates of the stator
N = the speed of the rotor in revolutions per second

The derivation of the constant $\pi/\sqrt{2}$ in this equation is based on two assumptions: 1. that the variation of capacitance as the device rotates is sinusoidal, and 2. that the impedance in the instrument circuit through which current flows is small when compared to the impedance between the rotor half and the stator half which it is facing. Note that the power to supply the output signal is supplied by the drive motor, and therefore this instrument draws no power from the room or device under test except for leakage current through 10^{14} ohms.

In practice this instrument can be calibrated by applying a known d-c potential to the stator of the rotating capacitor and observing the corresponding deflection of the milliammeter. Therefore, the exact wave shape of current output to the amplifier and its absolute magnitude are not of vital importance and the fact that the two foregoing assumptions are only approximately correct does not affect the over-all accuracy of the instrument.

It is obviously desirable to make I as large as possible to obtain the largest possible voltage across the load resistor. This requires the largest possible value for C as well as a high rotational speed for the motor. However, the larger C is made the slower the response to a change in voltage will be unless the effective resistance of the probe is decreased by increasing the radioactivity. In the instrument built by the authors, C was of the order of 50 micromicrofarads, the exact value depending on the spacing between the rotor and stator, which was varied during the course of experimentation. The resulting time constant with probe and cable was approximately 1 minute, which was satisfactory.

With fixed C and n, the output current and the voltage are directly proportional to the applied voltage, as shown in Fig. 6. The output voltage (rms reading of a 20-cycle-per-second signal) ranged from 0.6 to 1.0 per cent of the applied direct voltage, depending on the spacing between rotor and stator, so that a 10-volt potential on the stator supplied 60 to 100 millivolts at the brushes.

Early measurements were made with a motor operating at 3,400 rpm, and these measurements demonstrated the expected increase in output voltage with increased speed when compared with the 1,200-rpm motor. However, use of a frequency so close to line frequency demonstrated the problem of interference caused by stray field pickup, and a deliberate attempt was made to use a frequency quite different than line frequency, thus permitting compensation for any line signal pickup either in the probe or later in the circuit.

The Amplifier

The output of the rotating capacitor must be connected to a high-input-impedance vacuum-tube voltmeter so that the small current will develop as large a voltage as possible. The original calibration of the rotating capacitor was performed by using a commercially available instrument with an input resistance of 10 megohms. Fig. 7 shows the amplifier circuit that was constructed to work with this capacitor and examination of the circuit shows that it is a more or less conventional audio amplifier. A high input impedance was provided by using a free-grid cathode-follower input,[6] and a low output impedance to match standard 1-milliampere d-c recording instruments was obtained by again using a cathode follower in an electrometer-type output circuit. The amplifier was calibrated so that a direct voltage of 10 volts applied to the probe will drive a 1-milliampere meter to full scale. Analysis of this circuit shows that increased sensitivity would not be difficult to achieve if it were required since this amplifier requires approximately 100 millivolts of a 20-cycles-per-second signal for full-scale deflection of the meter. This sensitivity is satisfactory for presently contemplated measurements.

In meeting the fourth requirement cited earlier, namely, that the equipment must be capable of operation by people not especially trained in electrical measurement techniques, an attempt was made to keep the amplifier design as simple as possible. Feedback was therefore not used in this design and provision was made for simple calibration checks and recalibration if necessary (although these details are not shown in the schematic drawings). The possibility of drift in the amplifier was recognized and further experience may prove the necessity of utilizing a feedback type of design, in spite of the greater complication, although experience to date indicates that the amplifier design shown herein is sufficiently stable to allow operation by untrained people.

Applications

This instrument was designed for measurement and recording of atmospheric electrostatic potential as a means for determining the number of atmospheric ions present in a given environment. Although designed primarily for indoor use, it can be used outdoors if the probe is properly shielded. In this connection it is important to keep in mind that this instrument does not measure air ionization directly but measures instead the electrostatic potential from which the net ion density can be calculated. There appear to be two important limitations to be kept in mind when using this instrument for measuring of atmospheric electrostatic potential.

First, this instrument measures the total potential at a point regardless of the source of that potential. This means that stray electrostatic fields must be eliminated or must at least be known so that the instrument readings can be corrected. In many cases this presents no serious problem since, frequently, no stray electrostatic field is present; but in other situations careful examination of the readings is required. An example of this problem was observed by the authors when a recording was made with the unshielded probe outdoors, about 7 feet above ground, during a period of stormy weather. Potentials recorded varied from a few volts to several hundred volts, apparently in response to passage of clouds overhead.

Second, the use of electrostatic potential to determine ion density is complicated by the fact that only the difference in concentration between positive and negative ions can be measured and it is not possible to measure the total number of ions of each sign that may be present. Other types of instruments have been developed which measure ions of each polarity separately and one instrument is particularly well described by Fleming.[7] These instruments deplete the number of ions in the air and, therefore, in many circumstances do not give an accurate indication of the number of ions present in the absence of the instrument. This is in contrast with the instrument described in this paper which does not deplete the ions present in the atmosphere. Correlation of readings of these two types of instruments may be necessary in certain types of problems but where only the net ion density is required or where the information desired is atmospheric potential only, this instrument appears to have many possibilities of application.

Measurement of air ion density has been attempted for many years in conjunction with investigations of the effects of natural and artificial air ionization on humans and animals. Martin[8] summarizes much of the early investigation while discussing one means of artificially producing air ions. Skilling and Beckett[9] discuss other types of ion generators. Murphy[10] discusses the problem of unnatural ion balance from the viewpoint of the air-conditioning engineer. The instrument described herein is readily adaptable for use in testing the air ion density in ventilation ducts and rooms.

This instrument was devised as a means of measuring air ion density, since ion density can be calculated if the potential caused by the ions is known, as well as the room dimensions. Other possible applications include measurement of space charge where static accumulations present problems, and more general use as a d-c voltmeter having extremely high input impedance. The basic principles are well established and adaptable to many applications, while the resulting instrument is simple and reliable.

References

1. The Measurement of Electrostatic Potential Due to Net Ion Space Charge in Air, J. S. Carroll, S. B. Hammond. *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 712–15.

2. Production of Unipolar Air with Radium Isotopes, T. L. Martin, Jr. *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 771–76.

3. A Generating Voltmeter for Measurement of High Potentials, Paul Kirkpatrick, I. Miyake. *Review of Scientific Instruments*, New York, N. Y., vol. 3, Jan. 1932, pp. 1–8.

4. A Rotary Voltmeter, Paul Kirkpatrick. *Electrical Engineering*, vol. 51, Dec. 1932, pp. 863–65.

5. Further Developments of the Rotary Voltmetei, Paul Kirkpatrick. *Review of Scientific Instruments*, New York, N. Y., vol. 5, Jan. 1934, pp. 33–37.

6. Low-Noise Input Stage for Audio Preamplifier, James J. Noble, John K. Hilliard. *Electronics*, New York, N. Y., Feb. 1955, pp. 147–49.

7. Terrestrial Magnetism and Electricity (book), J. A. Fleming. Dover Publications, New York, N. Y., 1949, pp. 252–58.

8. Climate Control Through Ionization, T. L. Martin, Jr. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 254, Oct. 1952, pp. 267–80.

9. Control of Air Ion Density in Rooms, H. H. Skilling, J. C. Beckett. *Ibid.*, vol. 256, no. 5, Nov. 1953, pp. 423–34.

10. How Ion Density Affects Comfort, Howard Murphy. *Heating, Piping and Air Conditioning*, Chicago, Ill., Oct. 1954, pp. 120–25.

A Method of Solving Mathieu's Equation

J. J. SMITH
MEMBER AIEE

MATHIEU'S equation

$$\frac{d^2y}{dt^2}+(b-2q\cos 2t)y=0 \tag{1}$$

is gradually becoming a very important equation in physics and engineering. It apparently first appeared in a paper by Mathieu[1] on the vibration of an elliptic membrane. However, since that time it has made its appearance in many practical problems in the fields of mechanics, electricity and magnetism, sound, etc. Many such examples will be found in McLachlan's book on Mathieu functions.[2]

There seems to be a good reason why the interest in the solution of this equation is increasing. One of the differential equations most commonly encountered in physics and engineering is

$$\frac{d^2y}{dt^2}+by=0 \tag{2}$$

For instance, in mechanics this equation occurs when the system consists of a mass whose movement is restrained by a spring, in electrical problems it occurs when there is an inductance and capacitance in series. Thus this equation might be considered to apply to the simplest state of the system. In many practical cases, however, the actual system is not quite so simple and may either be more complicated by the addition of some driving force, or it may react on itself by a process which we may loosely term "feedback." In either case, some additional term must be added to the equation, and one useful form (there are, of course, others) is to add to equation 2 a term proportional to y of the form $2yq\cos 2t$ and then arrive at Mathieu's equation in equation 1.

To date, this equation has proved quite difficult to handle. Good descriptions of present-day methods will be found in Whittaker and Watson[3] and McLachlan.[2] In Whittaker and Watson several chapters are devoted to the hypergeometric functions before the chapter on Mathieu's functions. In the latter chapter they point out that Mathieu's functions are in a region beyond the hypergeometric functions, and these areas are as yet only very imperfectly explored.

Having recently published papers[4,5,6] showing the advantage of using Green's functions in the solution of equations of the hypergeometric type, it seemed worth

---

Paper 55-490, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted August 12, 1954; made available for printing May 12, 1955.

J. J. Smith is with the General Electric Company, Schenectady, N. Y.

The author acknowledges with pleasure the permission granted by the *Journal* of the Franklin Institute, and by W. L. Barrow, D. B. Smith, and F. W. Baumann, to reproduce Fig. 9 of their paper.

while to consider whether Green's functions might also be helpful in connection with Mathieu's equations. They are found helpful in giving solutions which are simple (though lengthy) and which show the performance of the circuit in a clear manner.

The method followed by Whittaker and Watson and by McLachlan is to assume a series solution for equation 1, to substitute it in the differential equation, and to try to evaluate the unknowns in the series in the usual manner. This leads in the general case to an infinite determinant whose zeros must be found to complete the solution. The determination of these zeros is not easy. They are important since they are related to the points at which the solution becomes unstable.

Instead of following this procedure the solution of equation 2 will be found when the circuit is initially in a state of rest and is excited at the time $t=0$ by an instantaneous impulse which is called the Green's function for the equation. This solution is analogous to 3-dimensional solutions in the theory of heat, etc., which have been called instantaneous Green's functions, and the similar solutions in the 1-dimensional case have also been called the Green's functions. From this Green's function for equation 2 it will be shown how the solution for any given exciting force applied to the right-hand side of equation 1 can be obtained. By this method the necessity of obtaining the zeros of an infinite determinant is avoided. Thus are found the points at which the circuit becomes unstable directly.

It is found, as will be shown that equation 2 is one whose solution is stable when it is excited by an impulse. When a driving force of $E \sin \omega t$ is added on the right-hand side of equation 2, the solution is stable when $\omega \neq b^{\frac{1}{2}}$. As is well known, it becomes unstable when $\omega = b^{\frac{1}{2}}$ corresponding to ordinary resonances.

In the case of equation 1, an equation whose solution is not stable when excited by an impulse, it will go into resonance at values of b equal to 1, 4, 9 . . . n^2 where n is an integer. When a driving force $E \sin \omega t$ is added on the right-hand side of equation 1, the solution becomes unstable for many other values of ω which will follow. In this case the terms "spontaneous resonance" or "autoresonance" are applied to the resonances obtained when the circuit is excited by an impulse and the term "forced resonance" has been applied to the resonances obtained when the circuit is excited by a driving force of the force $E \sin \omega t$.

Experimental data in the literature confirm these spontaneous and forced resonances and agree with the present-day theory of Mathieu functions as far as the latter goes.

Degenerate Form of Mathieu's Equation

To make clear the procedure to be used in solving Mathieu's equation, it will be applied first to the degenerate form when $q=0$. It then becomes the simple second-order differential equation with constant coefficients (changing b to m^2)

$$\frac{d^2y}{dt^2}+m^2y=0 \tag{3}$$

To obtain the Green's function for this equation (over an infinite interval), the right-hand side is put equal to the impulse function $H'(t,t_1)$ where $H'(t,t_1)$ is the derivative with respect to t of $H(t, t_1)$ and

$$H(t,t_1)=0 \quad \text{when} \quad t<t_1 \tag{4}$$

$$H(t,t_1)=1 \quad \text{when} \quad t>t_1$$

For those who do not like to differentiate such a discontinuous function, the more rigorous methods of Schwartz[7] are available. Schwartz's notation will not be used here since it is lengthy to write, but it can readily be used to verify any of the results to be given in this paper. A demonstration of this verification is given by Alger and the author.[8] Thus the equation to be solved is

$$\frac{d^2y}{dt^2}+m^2y=H'(t,0) \tag{5}$$

$H'(t,0)$ on the right is zero in the interval $(-\infty,\ 0-\epsilon)$. If y on the left is replaced by $yH(t,0)$, this ensures that the resulting solution is zero when $-\infty<t<0-\epsilon$. Thus this corresponds to the terminal condition that the circuit is initially at rest, giving the equation in the form

$$\frac{d^2y\,H(t,0)}{dt^2}+m^2y\,H(t,0)=H'(t,0) \tag{6}$$

For convenience in writing, contract $yH(t,0)$ to yH. To solve the equation, integrate it twice with respect to time from $0-\epsilon$ to t. Then all the integrals vanish at the lower limit since $H(0-\epsilon,0)=0$. Thus the resulting integral equation and its solution by successive approximations becomes

$$\begin{aligned} yH &= tH - m^2\int\int_{0-\epsilon}^{t} yH\,dt^2 \\ &= \left(t-\frac{m^2t^3}{3!}+\frac{m^4t^5}{5!}-\ldots\right)H+R \qquad (7) \\ &= \sin mt\,H/m \end{aligned}$$

Since equation 3 is a second-order differential equation, it must have another independent solution. One such solution can be obtained by putting the right-hand side equal to the impulse function of the second order $H''(t,0)$

$$\frac{d^2(yH)}{dt^2}+m^2y\,H=H''(t,0) \tag{8}$$

It is found by proceeding in the same manner that this equation has the solution

$$yH=\cos mt\,H \tag{9}$$

The physical meaning of such solutions is of interest. If we have a circuit consisting of an inductance and capacitance in series, and if a direct voltage EH be applied to it starting at the time $t=0$, the equation for the circuit is

$$L\frac{d(iH)}{dt}+\frac{1}{C}\int_0^t i\,H\,dt=EH \tag{10}$$

and, differentiated and rearranged, this becomes

$$\frac{d^2(iH)}{dt^2}+\frac{iH}{CL}=\frac{EH'}{L} \tag{11}$$

The solution is

$$iH=\frac{E\sin mtH}{mL} \tag{12}$$

where $m^2=1/LC$. Thus a current of frequency m will continue indefinitely in such a circuit (assuming there is no resistance) even though direct voltage is applied to the terminals. Similarly, when an impulse $H'(t,0)$ is applied to the inductance and capacitance in series, it is found that a current $\cos mt\,H$ persists indefinitely.

The solution corresponding to the application of the impulse voltage $H'(t,0)$ to an equation is called the Green's function for that equation. Thus for the equation

$$\frac{d^2(yH)}{dt^2}+m^2y\,H=H' \tag{13}$$

the Green's function is

$$G(t,0)=\sin mt\,H/m \tag{14}$$

From the Green's function the result can be found when a driving force given by $f(t)H$ is applied to the equation or

$$\frac{d^2(yH)}{dt^2}+m^2y\,H=f(t)H \tag{15}$$

by the well-known equation

$$yH=\int_0^t G(t-t_1)f(t_1)\,H\,dt_1 \tag{16}$$

For instance, let a driving force $E \sin \omega t$ be applied to equation 16. Then

$$yH = E\,m^{-1}\int_0^t \sin m(t-t_1)\sin \omega t_1\,H\,dt_1$$
$$= EH[\sin \omega t - \sin mt]/(\omega^2 - m^2) \qquad (17)$$

Here the first term corresponds to the current due to the impressed frequency $\sin \omega t$ and the second term depends on the resonant frequency of the circuit $\sin mt$. It should be noted that this latter current continues indefinitely.

Note that if $\omega = m$ the denominator approaches zero and y becomes infinite. For this case the value must be found of the current in the circuit by putting $\omega = m$ in the integral before performing the integration. Then

$$y = E\,\omega^{-1}\int_0^t \sin \omega(t-t_1)\sin \omega t_1\,H\,dt_1$$
$$= E\,tH\sin \omega t/2\omega \qquad (18)$$

Thus when $\omega = m$ the result contains the factor t and the current increases indefinitely with time. Thus for this value of ω the circuit is not stable. This is, of course, the well-known phenomenon of resonance. The complete solution is given for this simple case, since the similar solutions for Mathieu's equation will be left in the form of equation 17 and they can always be further evaluated by the process shown.

Instantaneous Green's Function for Mathieu's Equation

The so-called canonical form of Mathieu's equation is

$$\frac{d^2y}{dt^2} + (b - 2q\cos 2t)y = 0 \qquad (19)$$

Already it is found desirable in the degenerate equation to put $b = m^2$. It will also be found convenient to change the term $\cos 2t$ to $\cos 2at$. This is not essential but allows varying of the frequency with which the parameter of the circuit varies instead of referring all other frequencies to this as unity. In mathematical textbooks the independent variable generally used is z instead of t but this should cause no difficulty. As in the previous examples, the solution will be found when the circuit is initially at rest, and an impulse $H'(t,0)$ is applied at $t=0$. Thus, with these changes the equation becomes

$$\frac{d^2(yH)}{dt^2} + (m^2 - 2q\cos 2at)yH = H'(t,0) \qquad (20)$$

In this, rewritten

$$\frac{d^2(yH)}{dt^2} + m^2yH = H' + 2yq\cos 2at\,H \qquad (21)$$

the term $2\,yq\cos 2at\,H$ may be regarded as an exciting voltage like $E\sin \omega t$ in the foregoing, even though its magnitude depends on y. Therefore the Green's function is found as before for the equation

$$\frac{d^2(yH)}{dt^2} + m^2yH = H' \qquad (22)$$

which is $y = \sin mt\,H/m$ (23)

Then, using equation 16 gives the solution of equation 21 as

$$y = \sin mt\,H/m + 2q\,m^{-1}\int_0^t y\sin m(t-t_1)\cos 2at_1\,H\,dt_1$$
$$= \sin mt\,H/m + 2q\,m^{-2}\int_0^t \sin mt_1 \sin m(t-t_1)\cos 2at_1\,H\,dt + R_1 \qquad (24)$$

where

$$R_1 = 2q\,m^{-1}\int_0^t \sin m(t-t_1)\cos 2at_1 H\,dt_1 \times 2q\,m^{-1}\int_0^t y\sin m(t_1-t_2)\cos 2at_2\,dt_2 \qquad (25)$$

For a first approximation

$$y = \sin mt\,H/m - \frac{q}{2m}\left(\frac{\sin(m+2a)t}{2a(m+a)} - \frac{\sin mt}{m^2-a^2} - \frac{\sin(m-2a)t}{2a(m-a)}\right)H + R_1 \qquad (26)$$

Thus, with this method of solution it may already be seen that the circuit is not stable when $m = a$ if an impulse voltage is applied. This checks with the stability conditions already known.[9] It should be clear, however, that the regions of instability shown by McLachlan can come as close to the line $q = 0$ as needed but cannot actually touch it. This is because when $q = 0$

$$y = \sin mt\,H/m \qquad (27)$$

and there is no instability. It will be seen from the first approximation in the foregoing that the unstable term is

$$A = -2^{-1}q\left(\frac{\sin(m+2a)t}{2a(m+a)} - \frac{\sin mt}{m^2-a^2} - \frac{\sin(m-2a)t}{2a(m-a)}\right) \qquad (28)$$

giving

$$\underset{q\to 0}{Lt}\,A = 0 \qquad (29)$$

$$\underset{m\to a}{Lt}\,A = \infty \qquad (30)$$

Thus the unstable term does not converge uniformly as $q \to 0$ and $m \to a$ but takes the form 0/0. The same thing might be said in electrical terms, that there must be some coupling into the circuit of the form $2yq\cos 2at\,H$. Since this coupling produces resonance, the circuit is unstable no matter how small q is. If, however, $q = 0$ then the circuit is stable for an applied impulse voltage.

The second approximation may now be worked out, which gives

$$yH = \sin mt\,H/m - \frac{q}{2m}\left(\frac{\sin(m+2a)t}{2a(m+a)} - \frac{\sin mt}{m^2-a^2} - \frac{\sin(m-2a)t}{2a(m-a)}\right)H + \frac{q^2}{16m}\left[\frac{1}{2a(m+a)}\left(\frac{\sin(m+4a)t}{4a(m+2a)} + \frac{\sin mt}{4a(m+2a)} + \frac{t\cos mt}{m}\right) + \frac{1}{2a(m-a)} \times \left(\frac{\sin(m-4a)t}{4a(m-2a)} - \frac{\sin mt}{4a(m-2a)} - \frac{t\cos mt}{m}\right) - \frac{1}{m^2-a^2}\left(\frac{\sin(m+2a)t}{2a(m+a)} - \frac{\sin(m-2a)t}{2a(m-a)} - \frac{\sin mt}{m^2-a^2}\right)\right]H + R_2 \qquad (31)$$

With this second approximation it is seen that in addition to resonance at $m = a$ there is also resonance at $m = 2a$ when an impulse voltage is applied to the circuit. Similarly, it is evident that on taking the third and higher approximations that resonance will occur at the values of m corresponding to $m = 1a, 2a, 3a, 4a\ldots$, which are the values of m obtained in the usual theory of Mathieu's functions[9] (in slightly different notation).

As shown in the case of the degenerate equation, another solution of Mathieu's equation may be found by exciting it with a second-order impulse function. The equation to be solved is then

$$\frac{d^2(yH)}{dt^2} + (m^2 - 2q\cos 2at)yH = H''(t,0) \qquad (32)$$

Proceeding in a similar manner, the equation

$$\frac{d^2(yH)}{dt^2} + m^2yH = H''(t,0) \qquad (33)$$

has the solution

$$y = \cos mt\,H \qquad (34)$$

and hence the equation

$$\frac{d^2(yH)}{dt^2} + m^2yH = H''(t,0) + 2qy\cos 2at\,H \qquad (35)$$

has the solution

$$yH = \cos mt + 2q\,m^{-1}\int_0^t \sin m(t-t_1)\cos 2at_1\,yH\,dt_1$$
$$= \cos mt - \frac{qH}{2}\left(\frac{\cos(m+2a)t}{2a(m+a)} + \frac{\cos mt}{m^2-a^2} - \frac{\cos(m-2a)t}{2a(m-a)}\right) + R_3 \qquad (36)$$

Since an abbreviated name will be convenient for the solutions of these equations, and this should not conflict with the present terminology of $se(t,q)$ and $ce(t,q)$, the terms mas(t,q,m,a) and mas(t,q,m,a) will be used for lack of better terminology. Mas and mac are contractions of Mathieu sine and Mathieu cosine. In the usual terminology a suffix is used on *se* such as se_m and this is replaced in the present terminology by the m. The a is added for the present to make the results easier to use. However, it is not absolutely necessary since it is the ratio of m to a that is important, and thus it might be agreed to take $a=1$ for convenience in tabulation. Thus

$$\operatorname{mas}(t, q, m, a)H=\sin mt\, H/m-\frac{qH}{2m}\left(\frac{\sin(m+2a)t}{2a(m+a)}-\frac{\sin mt}{m^2-a^2}-\frac{\sin(m-2a)t}{2a(m-a)}\right)+R_1 \tag{37}$$

and

$$\operatorname{mac}(t, q, m, a)H=\cos mt\, H-\frac{qH}{2}\left(\frac{\cos(m+2a)t}{2a(m+a)}+\frac{\cos mt}{m^2-a^2}-\frac{\cos(m-2a)t}{2a(m-a)}\right)+R_2 \tag{38}$$

it may seem awkward to include the terms $\sin mt\, H/m$ and $\cos mt\, H$ in these solutions since these parts are due to the impulse functions. However, it may be safer to retain them for the present, since if there is no impulse excitation the circuit might not go into spontaneous unstable oscillation. It is evident that

$$\operatorname{mas}(t, q, m, a)=0 \text{ when } t=0 \tag{39}$$

However

$$\operatorname{mac}(t, q, m, a)\neq 0 \text{ when } t=0 \tag{40}$$

Thus we can write the general solution of Mathieu's equation for any initial conditions

$$y=A\operatorname{mas}(t, q, m, a)+B\operatorname{mac}(t, q, m, a) \tag{41}$$

and B can be obtained from the equation

$$(y)_{t=0}=B\operatorname{mac}(t, q, m, a)_{t=0} \tag{42}$$

Similarly, A is obtained from

$$\left(\frac{dy}{dt}\right)_{t=0}=A\left(\frac{d\operatorname{mas}(t, q, m, a)}{dt}\right)_{t=0} \tag{43}$$

In the foregoing what have been called spontaneous oscillations, are considered, i.e., oscillations which occur when the circuit is excited by d-c or is excited suddenly and the excitation then removed as, for instance, by an impulse function. What have been called forced oscillations, i.e., oscillations which occur when the circuit is excited by a sustained sinusoidal voltage, will now be considered.

Forced Oscillations of Mathieu's Equations

Let the applied voltage be $E\cos\omega t$ and let the current in the circuit be initially zero. Then Mathieu's equation takes the form

$$\frac{d^2(yH)}{dt^2}+(m^2-2q\cos 2at)yH=E\cos\omega t\, H \tag{44}$$

As before, this is rewritten

$$\frac{d^2(yH)}{dt^2}+m^2yH=2qy\cos 2at\, H+E\cos\omega t\, H \tag{45}$$

The Green's function, as before, is found for the left-hand side of this equation

$$y=\sin mt\, H/m \tag{46}$$

Then using equation 16 gives the solution of equation 45

$$y=m^{-1}\int_0^t \sin m(t-t_1)[2yq\cos 2at_1\, H+E\cos\omega t_1\, H]dt_1$$
$$=m^{-1}\omega^{-1}E\sin\omega t+m^{-1}\int_0^t 2yq\times\sin m(t-t_1)\cos 2at_1\, H\, dt_1+R_3 \tag{47}$$

The first approximation to the solution of this equation gives

$$y=m^{-1}\omega^{-1}E\sin\omega t+q\left[\frac{\sin(\omega+2a)t}{m^2-(2a+\omega)^2}+\frac{\sin(\omega-2a)t}{m^2-(2a-\omega)^2}-\sin mt\left(\frac{2a+\omega}{m^2-(2a+\omega)^2}-\frac{2a-\omega}{m^2-(2a-\omega)^2}\right)\right]+R_3 \tag{48}$$

Here it will be seen that resonance conditions occur when $m=2a+\omega$ or when $m=2a-\omega$. Further approximations will give additional values for m.

Another case occurs when the voltage of the exciting machine at no load is $E\cos\omega t$, but when it is connected to the circuit its terminal voltage is changed because of the current in the circuit. One assumption as to the form of the voltage might be that it now becomes

$$E\cos\omega t+y(a_1\cos\omega t+a_2\cos 2\omega t+a_3\cos 3\omega t+\ldots) \tag{49}$$

so that the change in voltage is proportional to y. The equation to be solved is now

$$\frac{d^2(yH)}{dt^2}+m^2yH=2qy\cos 2at\, H+E\cos\omega t\, H+y(a_1\cos\omega t+a_2\cos 2\omega t+a_3\cos 3\omega t+\ldots) \tag{50}$$

This equation is somewhat similar to Hill's[10] equation but is even more complicated.

As before, the Green's function is found for the left-hand side of this equation which is

$$y=\sin mt\, H/m \tag{51}$$

Then using equation 16 gives the solution of equation 50

$$y=m^{-1}\int_0^t \sin m(t-t_1)[2yq\cos 2at_1+E\cos\omega t_1+y(a_1\cos\omega t_1+a_2\cos 2\omega t_1+a_3\cos 3\omega t_1+\ldots)]H\, dt_1$$
$$=m^{-1}\omega^{-1}E\sin\omega t+m^{-1}\int_0^t \sin m(t-t_1)[2yq\cos 2at_1+y(a_1\cos\omega t_1+a_2\cos 2\omega t_1+a_3\cos 3\omega t_1+\ldots)]\times H\, dt_1+R_4 \tag{52}$$

The solution of this equation can be obtained by successive approximations as in the preceding examples. It is quite lengthy and will not be written here. It will be found that the resonance points are given by equations such as

$$m=xa\pm y\omega \tag{53}$$

where x and y are integers.

Comparison of Results with Experiment

The results given in the foregoing seem to check well with experiments which will be briefly referred to, particularly in connection with resonance conditions. To do this m, ω, q, and a are all taken to be real numbers in the following statements.

It is known that the equation

$$\frac{d^2(yH)}{dt^2}+m^2y\, H=H' \tag{54}$$

has only a stable solution; see equation 7.

It is known that the equation

$$\frac{d^2(yH)}{dt^2}+m^2y\, H=E\sin\omega t\, H \tag{55}$$

has a stable solution when $m\neq\omega$. When $m=\omega$ the solution is not stable since there is a resonance condition; see equations 17 and 18.

The equation

$$\frac{d^2(yH)}{dt^2}+(m^2-2q\cos 2at)yH=H' \tag{56}$$

was extensively studied experimentally by Barrow, *et al.*[11] They found that it went into a state of unstable spontaneous oscillation when $a=m$ in the notation given here. They report that the phenomenon was quite violent and also mention that mechanical couplings were broken when precautions were not taken. Oscillograms show the build-up of the current corresponding to resonance. They do not give any other frequency at which they looked for resonance. From the theory given in the foregoing it should also occur at $m=2a$, $3a\ldots$ and this is confirmed by the usual theory of Mathieu functions; see equation 26.

The forced oscillations of the same circuit were also studied by Barrow, *et al.* They point out that there was reaction

Table I. Frequencies for Various Values of y

x	$\frac{xa-m}{2\pi}$	$w/2\pi$ for				
		$y=1$	$y=2$	$y=3$	$y=4$	$y=5$
0	−60					
1	0	0	0	0	0	0
2	60	60	30	20	15	12
3	120	120	60	40	30	24
4	180	180	90	60	45	36
5	240	240	120	80	60	48
6	300	300	150	100	75	60
7	360	360	180	120	90	72
8	420	420	210	140	105	84
9	480	480	240	160	120	96

with the applied voltage. Hence the equation for their circuit may be taken as

$$\frac{d^2(yH)}{dt^2}+m^2y\,H=E\cos\omega t\,H+ 2qy\cos 2at\,H+y(b_1\cos\omega t+ b_2\cos 2\omega t+b_3\cos 3\omega t+\ldots)H \qquad (57)$$

It has been seen that the resonant frequencies for this circuit are given, equation 53, by

$$m=xa\pm y\omega \quad (x \text{ and } y \text{ integers}) \qquad (58)$$

In the case of their tests Barrow, *et al.* had

$$\frac{m}{2\pi}=60,\ \frac{a}{2\pi}=60$$

Thus data are obtained as shown in Table I. The frequencies they report are

$$\frac{\omega}{2\pi}=\ldots\ 30,\ 40,\ 50,\ 60,\ 80,\ 100,\ 120,\ 140,\ 160,\ 180\ \ldots$$

It will be noted that all these frequencies are given by the columns where $y=1, 2, 3$ except the frequency 50. This, of course, occurs at $y=6$ and $x=6$. Also, it will be noted in the column where $y=2$ there are two frequencies, 90 and 150, which were not observed. This might be because these are close to the observed frequencies 80 and 100 on either side of 90 and 140, and 160 on either side of 150. Or it also might be because the interaction term on the applied voltage was assumed to be of the form

$$y(a_1\cos\omega t+a_2\cos 2\omega t+a_3\cos 3\omega t+\ldots) \qquad (59)$$

There are no data available as to what it should be. If it is assumed that $a_2=0$, then the frequencies under the column $y=2$ do not appear and this is the column which contains the unobserved frequencies 90 and 150.

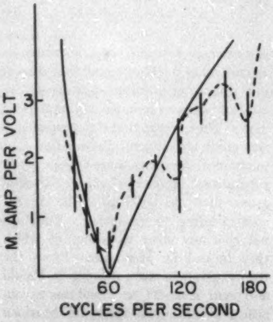


Fig. 1. Ordinary resonance curve (solid line) and Barrow, Smith, and Baumann's experimental curve (dashed line) for forced oscillations

In Fig. 9 of their article Barrow *et al.* give what they call the quasi resonance curve, and in the same figure they give the ordinary resonance curve. This is reproduced in Fig. 1 of this paper. The figure suggests that a rough approximation for such a circuit might be obtained by drawing the ordinary resonance curve for the circuit and then inserting on top of it resonances at the points given in the first three columns of Table I with the addition of a resonance at 50 and the omission of the resonances at 90 and 150 which require further investigation. This is illustrated in Fig. 2. It must be recollected that their ordinary resonance curve shows a very small amount of resistance, while in the actual tests enough resistance was inserted to avoid breakage. Considering all these differences, the results obtained by the suggested procedure give a surprisingly good approximation for a very complicated problem. Fig. 3 shows the spontaneous oscillations calculated in a similar manner.

Of course, a better approximation would be obtained by using the equations in this paper, which in turn might be modified to include the effect of resistance and thus to obtain an even better approximation. It seems, however, that so far the comparison with experiment shows good agreement and that additional studies in different cases by others will add to the knowledge of these types of circuits.

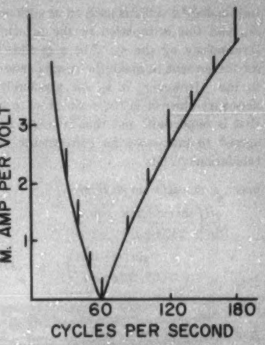


Fig. 2. Ordinary resonance curve with forced resonance peaks added at 30, 40, 50, 60, 80, 100, 120, 140, 160, 180 cycles, etc. Compare with dashed experimental curve in Fig. 1 as an approximation

Discussion

The solutions obtained for Mathieu's equations in the foregoing are similar to many other equations commonly encountered in mathematics. For instance

$$\frac{1}{\sin z}=\frac{1}{z}-2z\left(\frac{1}{z^2-\pi^2}-\frac{1}{z^2-4\pi^2}+\ldots+\frac{(-1)^{n-1}}{z^2-n^2\pi^2}+\ldots\right) \qquad (60)$$

Here the function has poles at $z=0, \pm\pi, \pm 2\pi,\ldots$

In a recent paper by the author[6] the equation for the vector potential in a rec-

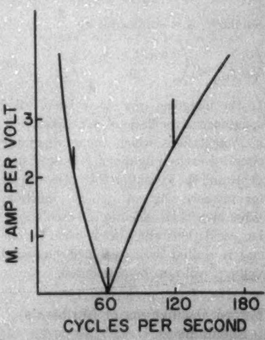


Fig. 3. Ordinary resonance curve with spontaneous resonance peaks added at 120 and 60 cycles, etc.

Vol.
Pt. 1
1955

tangular boxlike cavity came out to be

$$A_s = \frac{8Qu_zC}{abc}\sum_p\sum_q\sum_s \sin\alpha_p x \sin\alpha_p x_1 \times \sin\beta_q y \sin\beta_q y_1 \cos\gamma_s z \cos\gamma_s z_1 x \times \cos\omega(t-t_0)/(\omega^2-c^2\delta^2) \quad (61)$$

where $\alpha_p=p\pi/a$, $\beta_q=q\pi/b$, $\gamma_s=s\pi/d$; p, q and s take the values 0, 1, 2, 3... and $\delta^2=\alpha_p^2+\beta_q^2+\gamma_s^2$. For any given boxlike cavity if we apply a frequency ω such that $\omega^2-c^2\delta^2=0$, one of the terms becomes theoretically infinite and this term may be taken as the solution alone, as all other terms are negligible compared to it.

The solutions of Mathieu's equation given in equations 31, 36, and 48 have poles given by equation 53. Thus they tend to infinity at these points. Still more complex are the doubly periodic elliptic functions which have poles all over the z plane.

$$\wp(z)=\frac{1}{z^2}+\sum_{m,n}{}'\left(\frac{1}{(z-2m\omega_1-2n\omega_2)^2}-\frac{1}{(2m\omega_1+2n\omega_2)^2}\right) \quad (62)$$

where the ratio of ω_1 to ω_2 is not real.

Conclusion

The use of Green's functions as shown in the foregoing in the solutions of Mathieu's equation offers possibilities of solving nonlinear differential equations by splitting the process up into stages, at least in some cases. In the foregoing first the Green's function for the trigonometric equation was found. This Green's function was then used to form an integral equation from Mathieu's equation and the solution found by successive approximations. One of the advantages of this method is that, so to speak, it allows the differential equation to select the form of solution to which it can most readily adapt itself. In most of the standard approaches to this equation the usual procedure is to assume a solution and to try and verify that it satisfies the differential equation. Since the form of solution assumed is often not very suitable, this gives rise to considerable difficulty in showing that it satisfies the differential equation, and in many cases leads to the necessity of solving an infinite determinant. These difficulties do not appear in the method described in the foregoing, as no infinite determinants arise. The solutions obtained agree well with experiments. The results of experiments can be calculated approximately by finding the solution of the degenerate equation when $q=0$ and then adding the resonances to this curve at the points indicated by the solution as shown in the foregoing. Of course even better results can be obtained by using the equations given in the paper when this is necessary and worth while in view of the additional labor involved.

References

1. Memoire sur le mouvement vibratoire d'une membrane de forme Elliptique, **E. Mathieu.** *Journal de Mathématiques Pures et Appliques* (*Liouvilles Journal*), Paris, France, vol. 13, 1868, p. 137.

2. Theory and Application of Mathieu Functions (book), **N. W. McLachlan.** Oxford University Press, London, England, 1951. Contains an extensive series of 229 references.

3. Modern Analysis (book), **E. T. Whittaker, G. N. Watson.** Cambridge University Press, Cambridge, England, 1948.

4. Tables of Green's Functions, Fourier Series, and Impulse Functions for Rectangular Co-ordinate Systems, **J. J. Smith.** *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 22–30.

5. The Derivation of Vector Potential from Tables for Scalar Potential, **J. J. Smith.** *AIEE Transactions*, vol. 71, pt. I, Apr. 1952, pp. 169–74.

6. The Use of Instantaneous Point Sources or Green's Functions in Evaluating Electromagnetic Fields, **J. J. Smith.** *AIEE Transactions*, vol. 73, pt. I, Mar. 1954, pp. 82–88.

7. Theorie des Distributions (book), **L. Schwartz.** Hermann et Cie, Paris, France, 1950.

8. A Derivation of Heaviside's Operational Calculus Based on the Generalized Functions of Schwartz, **J. J. Smith, P. L. Alger.** *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 939–46.

9. See reference 2, p. 40, Fig. 8(A).

10. Mean Motion of the Lunar Perigee, **G. W. Hill.** Acta Mathematica, vol. 8, p. 1, 1886.

11. A Further Study of Oscillatory Circuits Having Periodically Varying Parameters, **W. L. Barrow, D. B. Smith, F. W. Baumann.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 221, 1936, pp. 403–16; 509–29.

Magnetic Flux Density Measurements in Rotating Machines

B. J. LEY
ASSOCIATE MEMBER AIEE

A. CHARMATZ
ASSOCIATE MEMBER AIEE

Synopsis: The Hall effect in a semiconductor is utilized to measure the flux density in the air gap of a rotating machine under transient conditions of starting and loading. This method offers several advantages over the conventional exploring-coil generator. Measurements of the flux density in a synchronous converter were made to demonstrate its applicability.

CONSIDERABLE information about the behavior of a rotating machine can be acquired from a knowledge of the magnetic-field intensity that exists in the air gap of the machine. One such method of determining the field is by means of an exploring-coil generator. The exploring-coil generator is essentially a standard d-c unit to which an exploring coil has been added. One side of the exploring coil lies in a conventional armature slot but the other side lies in a special groove in the armature shaft. The connections from the exploring coil are brought out separately to a set of collector rings at the opposite end of the armature shaft from the commutator. Thus when the armature revolves at a constant velocity, the flux linkages of the exploring coil vary as a function of the air-gap flux and hence a voltage proportional to the field strength is developed. Although there is much to be learned from a unit of this type, obviously a great deal is still to be desired. This paper illustrates another method that can be used to measure the flux density, in rotating machines, based on the Hall effect.

The Hall Effect

The Hall effect, discovered in 1879, occurs when a material carrying a current is subjected to a magnetic field which is perpendicular to the direction of the current density vector.[1,2] The effect is to

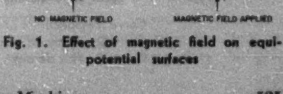


Fig. 1. Effect of magnetic field on equipotential surfaces

Paper 55-491, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted June 1, 1954; made available for printing May 11, 1955.

B. J. Ley is with New York University, New York, N. Y., and A. Charmatz is with the Oak Ridge School of Reactor Technology, Oak Ridge, Tenn.

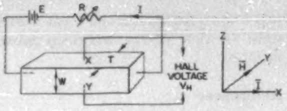


Fig. 2. Flux indicator based on Hall effect

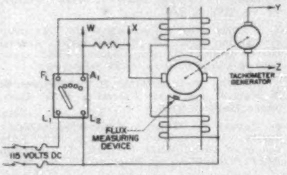


Fig. 3. D-c motor-starting connections

produce a transverse electric field that is proportional to the current density and the applied magnetic field.

For the rectangular piece of material illustrated in Fig. 1, the transverse electric field can be expressed mathematically as

$$\overline{\mathbf{E}}_H = -\nabla V_H$$

and $$|\nabla V_H| = \frac{V_H}{W} = R|\mathbf{i}||\mathbf{H}| = R\frac{I}{WT}H$$

or $$V_H = \frac{RIH}{T}$$

where

$\mathbf{E}_H$ = Hall electric field intensity
V_H = resulting Hall voltage
W = width of material
T = thickness of material
R = Hall constant of material
$\mathbf{H}$ = magnetic flux density
$\mathbf{i}$ = current density
I = magnitude of current flowing in material

Thus when the material is not under the influence of a magnetic field the currents flows longitudinally and there is no transverse electric field. However, in the presence of a magnetic field the current carriers experience a force and are deflected to the edges of the material. This change continues to build up until the electric field set up by this deflected charge exerts a force equal to the deflecting force of the magnetic field. Fig. 1 illustrates the equipotential surfaces in the material before and after the field has been applied.

Although the Hall effect is present in conductors, an equivalent sample using a semiconductor will result in a larger Hall voltage. The material used in making the following measurements was a piece of *p*-type germanium having a resistivity of approximately 6 ohm-centimeters and dimensions of approximately 1 by 1 millimeter (mm) by 1 centimeter. The current-carrying leads as well as the Hall voltage leads *X* and *Y* were soldered directly to the germanium sample. Although it is not possible to solder the Hall voltage leads *X* and *Y* exactly on the same equipotential surface, it is possible to buck out this small voltage under no-field conditions or to compensate in other ways.[3]

At this point the magnetic-field measuring possibilities of this device, Fig. 2, will be examined in contrast to the possibilities of the exploring coil mentioned earlier. First, it is essentially a point-measuring device and therefore capable of measuring the magnetic flux density at a given point. The exploring coil gives an indication that is proportional to the rate of charge of the integrated magnetic flux density variations over a given area as defined by the dimensions of the coil. Second, it indicates instantaneous magnetic fields and is therefore capable of measuring transient as well as d-c fields. The exploring coil gives an indication proportional to the rate of change of flux linkages and is therefore useful for only steady-state measurements. A third consideration must be given to the extremely small size of the semiconductor. Actually it can be made much smaller than the one used in these experiments. Several of these can therefore be distributed in a given machine to make simultaneous measurements at a number of different points. When the measurements have been completed, the units can easily be removed leaving the machine in its original condition.

Synchronous Converter Measurements

To observe the transient behavior in the starting of d-c as well as an a-c motor, a synchronous converter was chosen. The converter had the following rating: 2.5 kw, 110 volts d-c, 20 amperes d-c, 3-phase, 60 cycles per second (cps), 1,800 rpm. The converter was first disassembled so a *p*-type germanium specimen could be fastened by glyptal to one of the stationary pole faces of the machine (approximately located at center axially and one-third the way in). The four leads were brought out and the converter reassembled. The two current leads were connected to the battery and the resistor as shown in Fig. 2, and the resistor was adjusted to give a current of 10 milliamperes through the germanium sample. The two Hall voltage contacts were connected to a d-c amplifier of a Brush recording device. It might also be pointed out that for a better transient response the Hall voltage could be observed using a type-304 Dumont oscillograph (d-c amplifier input).

Fig. 3 illustrates the electric connections used for the d-c motor-starting measurements. Note that only two field poles are shown in this figure, while there were actually four. Terminals *WX* were connected to a second d-c amplifier and Brush recording device to record armature current. Terminals *YZ* were connected to a third d-c amplifier and Brush recording device to record the instan-

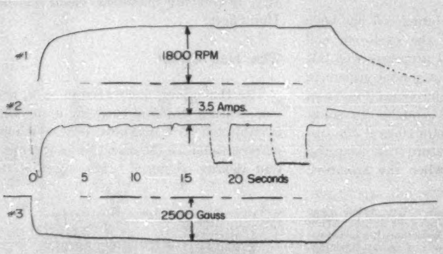


Fig. 4. D-c shunt motor-starting characteristics

Curve 1—Speed
Curve 2—Armature current
Curve 3—Flux density

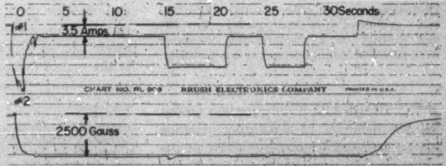


Fig. 5. Quick-starting d-c shunt motor characteristics

Curve 1—Armature current
Curve 2—Flux density

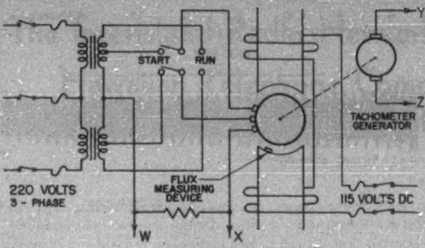


Fig. 6. A-c motor-starting connections

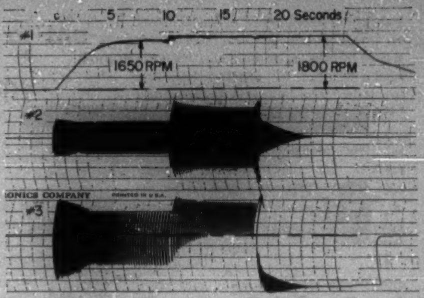


Fig. 7. A-c motor-starting characteristics

Curve 1—Speed
Curve 2—Line current
Curve 3—Flux density

taneous speed of the machine. Fig. 4 illustrates the resulting Brush recordings. The top curve indicates the speed (maximum deflection was 1,800 rpm), the middle curve indicates the armature current (3 amperes per 5 mm of deflection) as measured across a 20-ampere 50-millivolt ammeter shunt, and the bottom curve indicates the magnetic flux density (maximum deflection was approximately 2,500 gauss). The recording speed used for these curves was 5 mm per second.

The first part of these curves (approximately 15 seconds) is the starting period. The variations in the armature current were caused by manual changes in the starting-box resistance. For a starting period as long as this one, reductions in the starting resistance do not have an appreciable effect on the armature current and hence armature reaction is not very noticeable in the flux density curve. The shape of the magnetic flux density curve (as observed by other measurements) was found to be almost exactly the same as the field current build-up curve. For a shorter starting period (approximately 2.5 seconds, as shown in Fig. 5) armature reaction during starting is very much greater and its effect is more evident in the flux-density curve. It might also be pointed out that if the germanium sample had been located closer to the leading or lagging edge of the pole face, armature reaction effects would have been much more pronounced. At approximately 18 seconds after starting, Fig. 4, a balanced 3-phase resistance load was connected to the a-c side of the converter. Note the change in speed and the armature reaction effects in the flux density curve during the application as well as the removal of the load. Load was removed at 20 seconds and again applied at 24 seconds. At approximately 29.5 seconds after starting the line switch was opened. Note the reversal of the armature current due to the back electromotive force voltage and to the field current circuit path.

Fig. 6 shows the electric connections used for the a-c motor-starting measurements. Two transformers were connected in an open delta connection to supply 40 volts in starting and 80 volts for the running condition. A 100-ampere 50-millivolt shunt was used to record the line current (terminals *WX*) and the tachometer generator was again used to indicate the rotor speed (terminals *YZ*).

Fig. 7 shows the Brush recordings during starting conditions for a paper drive speed of 5 mm per second. The top curve again represented the speed of the rotor, the middle curve the 60-cps line current (using a 100-ampere 50-millivolt shunt), and the bottom curve the magnetic flux density measured at the same point on the stationary pole face as in the d-c starting measurements. The first 10 seconds show the effect of the application of half-starting voltage. Note that during this period the motor is running as an induction motor and levels off (in about 8 seconds) to a rotor speed of approximately 1,650 rpm (or a slip $=(1,800-1,650)/1,800=0.0833$). Note also the frequency of the flux density variations as measured on the stationary pole face level off from 60 to about 5 cps. The flux density curve indicates that the field poles were magnetized alternately north and south by the rotating a-c field and that they became more slowly magnetized as the armature approached synchronous speed. This can be observed by noting that there are approximately 5 cycles of flux density variation between 9 and 10 seconds. This might also be checked by computing the slip frequency at this time from the rotor speed as in the foregoing (for $f_s=sf=0.0833\times 60=5$ cps). At 10 seconds after starting full line voltage was applied. Note that hunting can be seen in both the speed curve and the line current curve. At about 18 seconds after starting, d-c field was applied. The surge in the line current as well as the flux density curve indicates that the polarity of the d-c field applied was opposite to the polarity to which the poles had become magnetized, because of the a-c field. All three curves illustrate that after the d-c field is applied, hunting gradually dies out and the machine runs at synchronous speed. The a-c line switch was opened at approximately 26 seconds after starting and the d-c field was removed approximately 3 seconds later. Note the effect in the slope of the speed curve when the d-c field was removed.

To study the a-c starting characteristics in more detail, the Brush recorders were operated at 25 mm per second. Fig. 8(A) illustrates the effect of the application of half-starting voltage. This expanded curve indicates that once the rotor speed begins to level off at around 1,600 rpm (approximately 5 seconds) uniform line current modulation can be seen in the 60-cps armature current curve. It can further be noted that this modulating frequency is equal to twice the slip frequency. The slip frequency can be determined from the frequency of the flux density variations as discussed. These variations can be explained by realizing that every time the rotating field passes one of the stationary poles (in a 4-pole machine) the reluctance decreases and this is reflected as a decrease in the line current. Thus, when the poles are demagnetized (as is originally the case), the line current variations show up as a twice slip frequency variation. It might also be noted that the minimum line current variations (maximum reluctance) occur almost at the same time (not at exactly the same time since the ger-

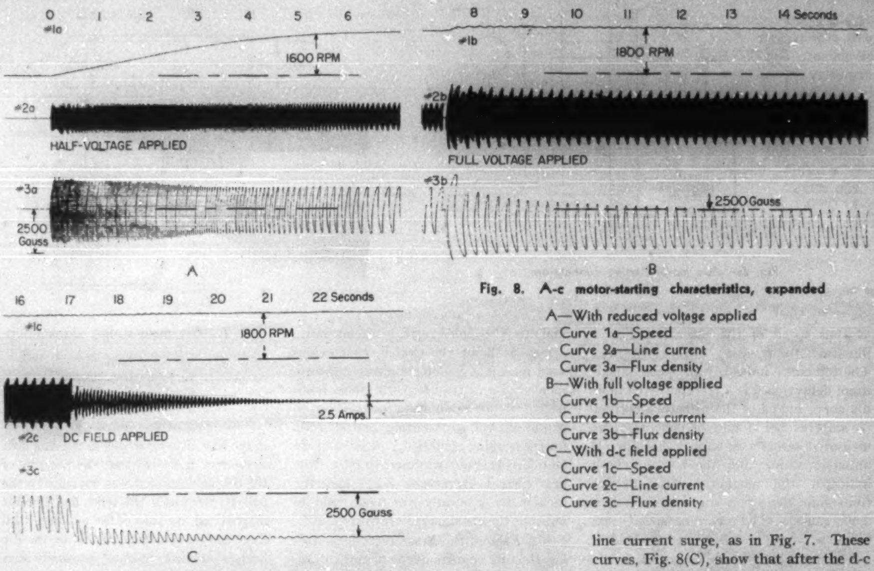


Fig. 8. A-c motor-starting characteristics, expanded

A—With reduced voltage applied
  Curve 1a—Speed
  Curve 2a—Line current
  Curve 3a—Flux density
B—With full voltage applied
  Curve 1b—Speed
  Curve 2b—Line current
  Curve 3b—Flux density
C—With d-c field applied
  Curve 1c—Speed
  Curve 2c—Line current
  Curve 3c—Flux density

manium sample is not located exactly at the pole center; it is about one-third the way in) as the positive maximums or negative maximums of the flux density variations.

Fig. 8(B) illustrates the effect of the application of full voltage on this expanded basis. Note here that after the initial application of full voltage the rotor speed increases to 1,800 rpm in approximately 1/10 second. It can further be noted from the rotor speed curve that hunting is taking place. It can also be seen that between 9 and 10 seconds after starting, this hunting frequency is approximately 5 cps. Since the rotor is now running at synchronous speed (1,800 rpm), the slip must be zero and therefore the flux density variations are not due to slip. A comparison of the rotor speed curve with the flux density curve shows that the variations in the rotor speed curve are the same as the variations in the flux density curve. The flux density variations are therefore now due to rotor hunting. Observe that after the initial application of full voltage (7.5 seconds) there is a gradual shift in the envelope of the flux density curve away from the original zero position, toward the bottom of the chart strip. This d-c magnetization is due to synchronization.

The variations in the line current curve Fig. 8(B), are now due to hunting and not slip. Note here that after the initial application of full voltage the alternate peaks in the line current variations gradually die out at an exponential rate leaving the resulting variations equal to exactly the hunting frequency. This condition can be explained by noting that previous to the application of full voltage the flux density curve indicates little d-c magnetization of the field poles, but when the full voltage was applied the flux density curve indicates the poles were becoming magnetized slowly at some exponential rate. This exponential rate of drift of the flux density curve due to the slow d-c magnetization is exactly the same rate at which the alternate line current variations are decreasing. Thus the periodic swinging of the rotor about a decreasing power angle causes the alternate peaks to be reduced.

Fig. 8(C) illustrates the effect of applying the d-c field. Note in this instance the polarity to which the poles had become magnetized; since the direct voltage was of the polarity to aid the field, there was no line current surge, as in Fig. 7. These curves, Fig. 8(C), show that after the d-c field was applied, the hunting gradually dies out and the machine runs as a synchronous motor.

Conclusion

A relatively simple flux density indicator, for the measurement of the magnetic field in a rotating machine, can be made from a semiconductor, based on the Hall effect. Its extremely small size would permit a number of these devices to be placed at different points in the air gap of the machine. Several different simultaneous point measurements could therefore be made on the rotating machine under transient conditions. The authors feel that such a device would be of material aid from an educational as well as an experimental standpoint.

References

1. HALL EFFECT, O. Lindberg. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 40, Nov. 1952, pp. 1414–19.

2. ELECTRONS AND HOLES IN SEMICONDUCTORS (book), W. Shockley. D. Van Nostrand Company, New York, N. Y., 1950, pp. 4–15; 204–05; 209–17.

3. A MAGNETIC FIELD STRENGTH METER EMPLOYING THE HALL EFFECT IN GERMANIUM, G. L. Pearson. *Review of Scientific Instruments*, New York, N. Y., vol. 19, 1948, pp. 263–65.

The Interisland Radio Network of the Hawaiian Telephone Company

W. B. FARINON
NONMEMBER AIEE

TEN years of planning and 2½ years of actual installation work recently came to completion when the Hawaiian Telephone Company placed new radio equipment in operation on the interisland circuit between the islands of Oahu and Lanai. This completed the entire network which now provides service to six major islands of the Hawaiian chain.

The Hawaiian Islands and Their Telephone Company

The Hawaiian Islands include Kauai, Oahu, Maui, Molokai, Lanai, and the largest island, Hawaii. As shown in Fig. 1, these islands are located in an area approximately 2,000 miles from the coast of California with relative distances between islands as shown on the chart. The interisland radio network between these islands is probably one of the most extensive installations of its kind in the world with relatively large circuit groups transmitted over long open-water paths between islands. Its development has required consideration of many unusual and unique engineering and operation problems which are not ordinarily encountered by communication carriers common on the mainland.

The end of hostilities in the Pacific and the beginning of expansion in the Hawaiian Islands brought about the need for greater communication facilities, both interisland and transpacific. At the end of World War II, the company had 13 voice channels interconnecting the six main islands. All radio circuits, some of which were only single-channel, operated in the 30- to 50-megacycle (mc) range. As the demand for circuits began to grow, additional carrier-derived channels were added to the previously installed radio circuits. In addition, four new radio links with somewhat larger channel capacity were operated in the 152- to 162-mc range. By 1950, interisland facilities had been increased to 35 3-kc channels, with 27 of these used for voice, four used for dial, and the remaining four used for telegraph. Even though the number of interisland voice circuits had been more than doubled, it was evident by 1950 that the existing radio network would soon prove totally inadequate to meet the traffic demands forecast for the next 10 or 20 years.

In 1944, when the real planning for the new radio network began, the total population of the Hawaiian Islands was 464,250. There were 55,271 telephones in use. The population is now 492,000 and there are 132,822 telephones. Because the islands served by the Hawaiian Telephone Company are very much interdependent from an economic standpoint, interisland communication facilities are particularly important.

Most of the business activity for the Territory is concentrated in Honolulu, on the island of Oahu, with constant and reliable contact with offices on the other islands forming a practical necessity of everyday life. Here life, in many ways, is quite similar to patterns of life in any of the mainland states. A subscriber of the Hawaiian Telephone Company in Hilo, on the island of Hawaii, is just as likely to place a call to New York or Washington, D. C., as is a subscriber of any mainland company residing in any similar city. Furthermore, the Hilo subscriber expects the same type of service as is available any place else in the United States.

The islands are the summits of a series of vast mostly submerged volcanic mountain peaks rising from great depths up to 30,000 feet beneath the sea. On the island of Hawaii volcanic activity still exists and two mountains rise to the height of over 13,000 feet above sea level. One is the active volcano Maunaloa and the other Maunakea which is extinct. The other islands are older and have been subjected to long periods of erosion. Most of the land is rugged with huge peaks and ridges separated by deep fertile valleys. Some topographical problems involved in providing communications can be seen in the aerial view of Fig. 2.

Development of Interisland Radio System

Interisland radiotelephone service has been provided on a regular basis in the islands since 1931. Transpacific service was installed on a commercial basis between Honolulu and Oakland in the same year. However, the first attempts at electrical communication between islands date back to 1889. In that year a local company was organized under what was then the Kingdom of Hawaii, and a cable was laid from the island of Oahu to the island of Molokai. Unfortunately, the channel between Oahu and Molokai proved too deep, too rough, and too rocky, and the cable parted soon after it was installed. After one unsuccessful attempt to splice the cable, the entire project was abandoned.

In 1889, a contract was signed with Marconi to supply wireless equipment to connect five of the islands. Many difficulties were experienced before the first radio telegraph messages were sent commercially in March of 1901. Between 1901 and 1907, the original telegraph company went through many financial crises before it finally achieved a sound footery. The interisland telegraphy company also engaged in ship-to-shore communication, with a station completed in 1908. On October 15 of that year, when the station was completed, night communication was established with San Francisco, 2,100 miles away. This was the first direct radio communication between Hawaii and the mainland, and was a record for long-distance working at that time. The inter-island system soon became part of the Mutual Telephone Company (now Hawaiian Telephone Company) and the radio telegraph stations formed the basis for interisland radio telephone communication in 1931.

Between 1931 and the end of World War II, the radiotelephone network grew slowly. Then, as the war drew to a close, the Hawaiian Telephone Company began to plan major changes in its interisland service. Extensive studies and forecasts were made to determine the best methods to use. One of the first plans considered was the use of a central repeater point on

---

Paper 55-534, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 25, 1955; made available for printing May 6, 1955.

W. B. FARINON is with the Hawaiian Telephone Company, Honolulu, Hawaii.

The Management and Engineering Staff of the Hawaiian Telephone Company have greatly appreciated assistance given by other organizations during the planning and installation of the new interisland network. Detailed assistance on system engineering was obtained from engineers of the Lenkurt Electric Company and the Automatic Electric Sales Corporation, among others. In addition, advice was sought and freely given by operating personnel such as B. R. Tupper, Northwest Telephone Company, F. W. Ryan, American Telephone & Telegraph Company, and many others.

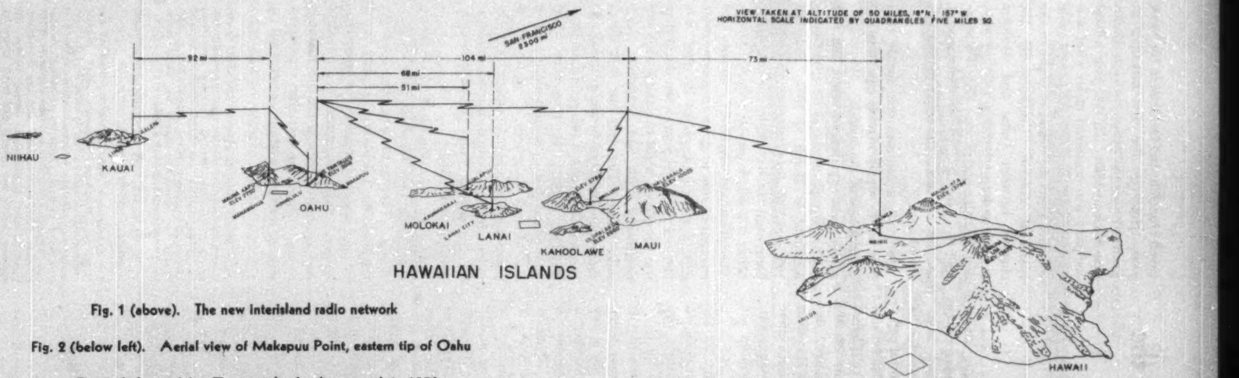

VIEW TAKEN AT ALTITUDE OF 50 MILES, 18° N, 157° W
HORIZONTAL SCALE INDICATED BY QUADRANGLES FIVE MILES SQ.
SAN FRANCISCO 2300 mi
92 mi
104 mi
68 mi
51 mi
73 mi
NIIHAU
KAUAI
KALEPA
LIHUE
MAUNA KAPU ELEV 1750
MANAWAHUA
HONOLULU
OAHU
MAKAPUU
MOLOKAI
KAUNAKAKAI
LANAI CITY
LANAI
KAHOOLAWE
ELEV 5788
WAILUKU
HALEAKALA ELEV 10025
ULUPALAKUA ELEV 2660
MAUI
MAUNA KEA ELEV 13784
WAIKII
HILO
KAILUA
HAWAII
HAWAIIAN ISLANDS

Fig. 1 (above). The new interisland radio network

Fig. 2 (below left). Aerial view of Makapuu Point, eastern tip of Oahu

Fig. 3 (below right). The interisland radio network in 1950

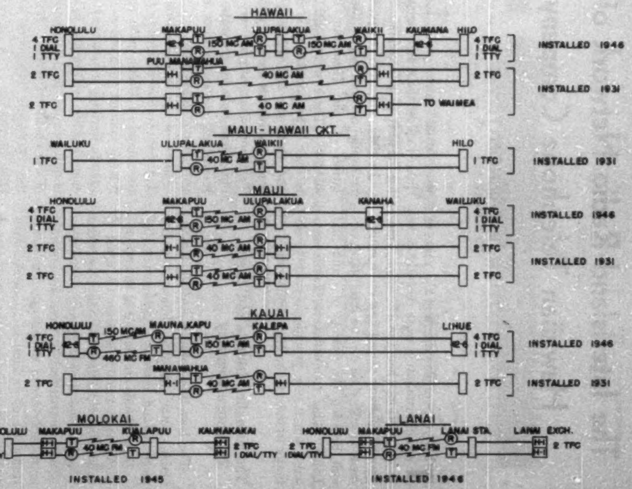

HAWAII
HONOLULU
MAKAPUU
ULUPALAKUA
WAIKII
KAUMANA
HILO
4 TFC
1 DIAL
1 TTY
150 MC AM
INSTALLED 1946
PUU MANAWAHUA
2 TFC
40 MC AM
TO WAIMEA
INSTALLED 1931
MAUI - HAWAII CKT.
WAILUKU
ULUPALAKUA
WAIKII
HILO
1 TFC
40 MC AM
INSTALLED 1931
MAUI
HONOLULU
MAKAPUU
ULUPALAKUA
KANAHA
WAILUKU
4 TFC
1 DIAL
1 TTY
150 MC AM
INSTALLED 1946
2 TFC
40 MC AM
INSTALLED 1931
KAUAI
HONOLULU
MAUNA KAPU
KALEPA
LIHUE
4 TFC
1 DIAL
1 TTY
150 MC AM
460 MC FM
150 MC AM
INSTALLED 1946
MANAWAHUA
2 TFC
40 MC AM
INSTALLED 1931
MOLOKAI
HONOLULU
MAKAPUU
KUALAPUU
KAUNAKAKAI
2 TFC
1 DIAL/TTY
40 MC FM
INSTALLED 1945
LANAI
HONOLULU
MAKAPUU
LANAI STA.
LANAI EXCH.
2 TFC
1 DIAL/TTY
40 MC FM
INSTALLED 1946
T
R
H-I

Fig. 4 (left). Old and new radio stations at Ulupalakua, Maui. Old station at bottom of hill included housing for maintenance personnel

Fig. 5 Transmitting building at Tantalus provides facilities for circuits from Oahu to Hawaii, Maui, Lanai, and Molokai

the island of Maui with separate radio links to each of the other islands.

The summit of Haleakala, at an elevation of 10,000 feet was believed to provide line-of-sight paths that would reach all islands except Kauai which could be reached by use of a repeater station on Oahu. The plan was to use 6,000-mc equipment, but further investigation proved this to be impractical. Later the entire central repeater was abandoned in favor of a more straightforward expansion using radio circuits radiating outward from the Island of Oahu on which the main city of Honolulu is situated.

All of the very high frequencies in use prior to the beginning of the expansion program were, with minor exceptions, allocated primarily for services other than common carrier. Interference to or from other services was a constant problem. The transmission characteristics of frequencies in the 30- to 50-mc band were such that the interisland radio networks had on some occasions caused interference to stations on the mainland as well as in the islands where some of the public safety radio services with a priority on the frequency band could prevent use of the interisland frequencies for long periods of time. In one instance, salvage operations were being carried out on a freighter which went aground off the coast of Kauai. During the 3 weeks that the salvage operations were in progress, the 30- to 50-mc Oahu-Kauai circuit was forced to suspend operations for long intervals. Excessive fading also proved to be a problem on one of the 30- to 50-mc circuits operated over a 180-mile beyond-line-of-sight path between Oahu and Hawaii.

In addition to the interference problems involved in the continued use of 30 to 50 mc, much of the company's equipment for operation in that band had been in service as early as 1931. This equipment was obsolete, both mechanically and electrically, and continuation of reliable service to subscribers required that such equipment be replaced as soon as possible. Continued use of the 152- to 162-mc band also presented a number of complications. Although transmission characteristics in this band were essentially satisfactory, other users or potential users in the islands had priority in their use. Furthermore, the band was not sufficiently broad to provide all of the interisland circuits needed.

At the time expansion plans were being formulated, the only frequencies allocated exclusively for common carrier use were above 3,700 mc. Although such frequencies were adequate for the overland radio circuits required in continental United States, they were not suitable for the overwater interisland routes.

Since the previously used very-high-frequency bands were inadequate for an expanded interisland network, and since the common carrier bands normally available were unsuitable for use, an intensive survey was conducted by company personnel in 1950 to select and to obtain a suitable band of frequencies. To confirm available data on propagation characteristics at various frequencies, a series of tests was made on 150, 460, and 1,900 mc. These tests showed that propagation at 1,900 mc was completely inadequate, that 460 mc could have been used for one of the links but, if adequate circuits to all islands were to be obtained, most of the circuits would have to operate in a frequency band below 250 mc. Only frequencies above 50 mc could be considered because of the possibility of causing interference to mainland stations. Further study showed that, to obtain the overwater circuits then considered necessary to meet the ultimate objectives of the interisland network, two 10-mc bands, separated by a band guard of about 10 mc, would be required.

Because the frequencies between 50 and 250 mc were allocated primarily to television, frequency modulation, and government services, no frequency bands suitable for the interisland circuits were readily available. The only frequencies which might practically be used for Hawaiian interisland circuits, taking into consideration all of the many factors involved, were television channels 5 and 6 (76 to 88 mc) and frequency-modulated channels 251 to 300 (98 to 108 mc).

At the time of this study it was found that television channels 5 and 6 were not allocated for use in the Hawaiian Islands, and that these channels would, in all probability, not be required to provide adequate television service to the area. Also, at that time, no person had sought permission to construct even one frequency-modulated station in the islands. It was highly unlikely that full utilization could ever be made of the 100 assignable frequency-modulated channels since the entire islands were then being adequately served by only 12 amplitude-modulated broadcast stations.

On the basis of the considerations briefly outlined in the foregoing, the Hawaiian Telephone Company presented a petition to the Federal Communication Commission requesting a special allocation of the 76- to 88-mc and the 98- to 108-mc bands for common carrier interisland service. When this petition was granted in 1952, a full-scale effort was begun to implement the entire expansion and modernization program as quickly and efficiently as possible. In addition to the special allocation of two very high frequency bands for interisland radio links, permission was also obtained to operate certain intraisland overland radio links in the 900-mc band.

Expansion Problem and Solution

Interisland telephone service in 1950 was provided by the radio network shown in Fig. 3. This consisted of 11 radio circuits, each providing from one to six 3-kc channels, some of which were used for dial and telegraph requirements. Two of the radio circuits included repeaters. Only one circuit, that between Maui and Hawaii, did not carry traffic terminating on one end in Honolulu.

The 1950 network had many disadvantages which affected performance and often made operation difficult. All but two of the stations were operated on an attended basis and, in most cases, were located some distance from the connected exchange and nearest population center. Consequently, it was necessary to include many special provisions such as housing for personnel at each site. The large number of systems had also necessitated the installation of large quantities of equipment at Ulupalakua on Maui with consequent problems of mutual interaction and interference. Since the station at Makapuu Point was 17 miles from the Honolulu toll office, voice frequency amplifiers and hybrids were required for the termination of the circuits in Honolulu. Circuits were extended on a 4-wire basis through interconnecting cable between Makapuu and Honolulu. Finally, because the network had grown over a period of almost 20 years, many different types of equipment were in use, and many locally devised methods had been employed to obtain the additional circuits added during these years.

DESCRIPTION OF THE 1950 NETWORK

Three radio circuits provided facilities between Oahu and Hawaii before the expansion program was initiated. Two of these were nonrepeatered 40-mc amplitude-modulated systems, one providing two voice circuits to Waimea, and one providing two voice circuits to Hilo. The third circuit, providing six 3-kc channels (four voice, one dial, one telegraph) to Hilo, operated in the 150-mc range and utilized a repeater at Ulupalakua, Maui.

The Oahu terminal for the 150-mc circuit was at Makapuu Point on the southeast tip of the island. The 40-mc circuits terminated at Puu Manawahua on the slopes of the Western mountain range. All three terminals on the island of Hawaii were at Waikii, an isolated site almost 50 miles from the principal city of Hilo.

In the 1950 network, service between Maui and Oahu was provided by an arrangement very similar to that between Hawaii and Oahu with one 150-mc and two 40-mc circuits. All radio links terminated at Makapuu Point on Oahu and at Ulupalakua, Maui, a site on the Western slope of Mt. Haleakala. An open-wire line was used for transmission of the multiplexed channels between Ulupalakua and the exchange at Wailuku. The Ulupalakua site is shown in Fig. 4. The large group of buildings shown in the lower part of the picture is the old station used prior to 1954. It includes housing for personnel and is typical of the stations in use at that time. Only one building, that on the left in the upper part of Fig. 4, is now in use.

Service between Oahu and Kauai was provided by one 40-mc circuit with two channels and one 150-mc circuit with six channels. On Oahu the 40-mc link terminated at Manawahua, and the 150-mc link terminated at Mauna Kapu. An unusual repeater section was used between Mauna Kapu and Honolulu. It utilized 150-mc amplitude-modulated equipment in the outward direction from Honolulu. However, to avoid interference which was becoming a problem in the overcrowded 150-mc band, 460-mc frequency-modulated equipment was used in the other direction. Kauai radio ter-

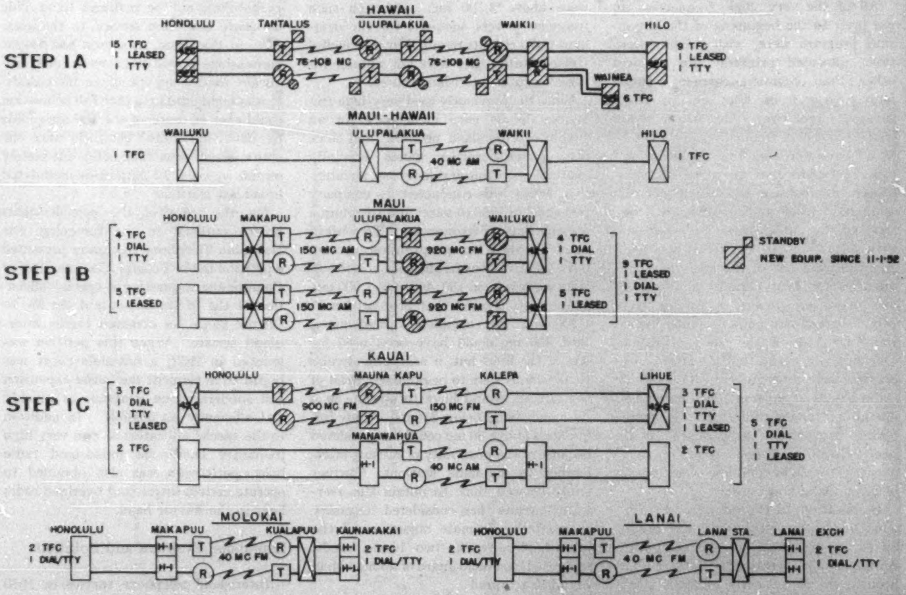

HAWAII
HONOLULU
TANTALUS
ULUPALAKUA
WAIKII
HILO
STEP 1A
15 TFC
1 LEASED
1 TTY
76-108 MC
76-108 MC
WAIMEA
6 TFC
9 TFC
1 LEASED
1 TTY
MAUI - HAWAII
WAILUKU
ULUPALAKUA
WAIKII
HILO
1 TFC
40 MC AM
1 TFC
MAUI
HONOLULU
MAKAPUU
ULUPALAKUA
WAILUKU
STEP 1B
4 TFC
1 DIAL
1 TTY
150 MC AM
920 MC FM
4 TFC
1 DIAL
1 TTY
5 TFC
1 LEASED
150 MC AM
920 MC FM
5 TFC
1 LEASED
9 TFC
1 LEASED
1 DIAL
1 TTY
STANDBY
NEW EQUIP. SINCE 11-1-52
KAUAI
HONOLULU
MAUNA KAPU
KALEPA
LIHUE
STEP 1C
3 TFC
1 DIAL
1 TTY
1 LEASED
900 MC FM
150 MC FM
MANAWAHUA
2 TFC
H-1
40 MC AM
H-1
3 TFC
1 DIAL
1 TTY
1 LEASED
2 TFC
5 TFC
1 DIAL
1 TTY
1 LEASED
MOLOKAI
HONOLULU
MAKAPUU
KUALAPUU
KAUNAKAKAI
2 TFC
1 DIAL/TTY
H-1
40 MC FM
2 TFC
1 DIAL/TTY
LANAI
HONOLULU
MAKAPUU
LANAI STA.
LANAI EXCH
2 TFC
1 DIAL/TTY
40 MC FM
2 TFC
1 DIAL/TTY

Fig. 6. Interisland radio network, October 15, 1953

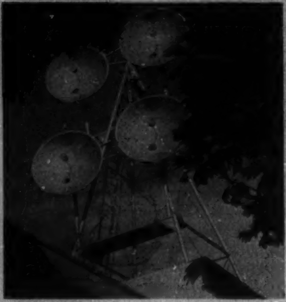


Fig. 7. Antennas on the roof of the Wailuku exchange building for the 900-mc terminals of the Oahu-Maui radio circuit

minals for both circuits were located on Kalepa Ridge only a few miles from the Lihue exchange on Kauai. The other three radio circuits in operation in 1950 provided service from Honolulu to the town of Kaunakakai, Molokai, from Honolulu to Lanai City, Lanai, and from Wailuku, Maui, to Hilo, Hawaii.

Expansion Program

Allocation of the 76- to 88- and 98- to 108-mc bands permitted the company to plan a radio network which would eliminate most of the previous operating problems while providing better and more reliable service. By using appropriate types of equipment, a permanent network could be installed which would be capable, without additional radio circuits, of providing gradually increased channel capacity to each of the islands as traffic increased over a considerable future period. By using new types of broadband radio and new types of channelizing equipment, the larger circuit groups needed could be obtained. And through elimination of some of the old radio sites, with operation of the remotely located sites on an unattended basis, many of the personnel problems could be alleviated. Accordingly, a 4-step expansion program was established which, over a period of 2 to 3 years, would replace all of the previously existing radio circuits with new wide band frequency-modulated to which additional multiplexing equipment could be added as needed to meet future circuit requirements. The program was initiated late in 1952.

One of the first requirements for the program was to establish a new radio site on Oahu which could serve as a terminal for all of the systems except those to Kauai. An ideal site was found on Tantalus, a 2,013-foot elevation about 4 miles from the Honolulu exchange. Since it was desirable to locate all of the channelizing equipment in the Honolulu Exchange building, two cables were installed to Tantalus with one used for transmitting and one for receiving the carrier frequencies of four independent radio circuits. The Tantalus location provides line-of-sight paths to Maui, Molokai, and Lanai. Loss from Tantalus to the Ulupalakua station on Maui is 8 decibels less than from the old station at Makapuu. Fig. 5 shows the transmitting building at Tantalus. Because the Western Oahu mountain range blocked the path from Tantalus to Kauai, the Mauna Kapu repeater site was retained for the Kauai circuits.

The equipment installed during the first step of the expansion program is indicated in Fig. 6. Five new radio links were included in this step. Two of these links comprise an Oahu-Hawaii system repeatered at Maui which replaced the three previous systems connecting these two islands. Three carrier systems, one Lenkurt type *42C* and two Lenkurt type *32E* were specially modified to provide 12 voice channels to Hilo and six voice channels to Waimea over the single radio circuit. An example of the manner in which equipment was shifted to meet urgent channel requirements can be seen in the

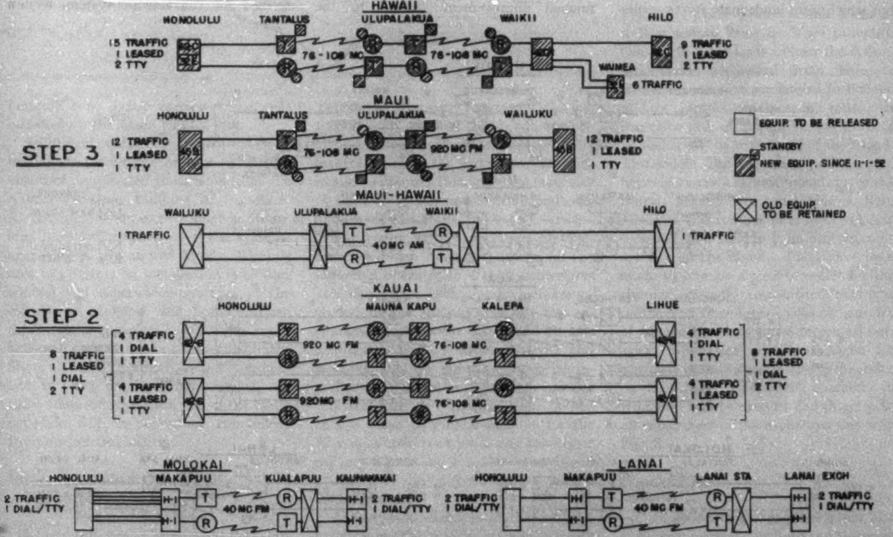

HAWAII
HONOLULU
TANTALUS
ULUPALAKUA
WAIKII
HILO
15 TRAFFIC
1 LEASED
2 TTY
76-108 MC
76-108 MC
WAIMEA
9 TRAFFIC
1 LEASED
2 TTY
6 TRAFFIC
MAUI
HONOLULU
TANTALUS
ULUPALAKUA
WAILUKU
STEP 3
12 TRAFFIC
1 LEASED
1 TTY
76-108 MC
920 MC FM
12 TRAFFIC
1 LEASED
1 TTY
EQUIP. TO BE RELEASED
STANDBY
NEW EQUIP. SINCE 11-1-52
OLD EQUIP. TO BE RETAINED
MAUI-HAWAII
WAILUKU
ULUPALAKUA
WAIKII
HILO
1 TRAFFIC
T
R
40MC AM
1 TRAFFIC
KAUAI
STEP 2
HONOLULU
MAUNA KAPU
KALEPA
LIHUE
8 TRAFFIC
1 LEASED
1 DIAL
2 TTY
4 TRAFFIC
1 DIAL
1 TTY
4 TRAFFIC
1 LEASED
1 TTY
920 MC FM
76-108 MC
4 TRAFFIC
1 DIAL
1 TTY
4 TRAFFIC
1 LEASED
1 TTY
8 TRAFFIC
1 LEASED
1 DIAL
2 TTY
MOLOKAI
HONOLULU
MAKAPUU
KUALAPUU
KAUNAKAKAI
2 TRAFFIC
1 DIAL/TTY
40 MC FM
2 TRAFFIC
1 DIAL/TTY
LANAI
HONOLULU
MAKAPUU
LANAI STA
LANAI EXCH
2 TRAFFIC
1 DIAL/TTY
40 MC FM
2 TRAFFIC
1 DIAL/TTY

Fig. 8. Interisland radio network, August 1954

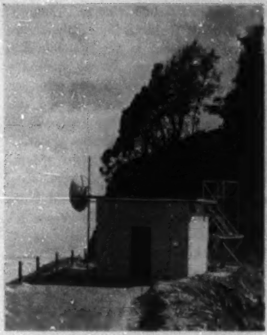


Fig. 9. Receiver building at Mauna Kapu

treatment of the 150-mc equipment previously installed between Oahu and Hilo. This equipment was placed in service between Oahu and Maui to replace the previously installed 40-mc systems between the two islands as a stop-gap measure until step 2 of the program could be completed.

The other three radio links installed in step 1 were 900-mc equipment for intra-island use. Two were used as repeater sections between Ulupalakua and Wailuku to replace wire lines formerly used, but which were inadequate for transmission of the carrier-frequency bands used on the radio links to Ulupalakua. The other 900-mc link was placed in service between the Honolulu exchange building and the Mauna Kapu radio site for the Kauai systems. This replaced the previously used 150- to 460-mc radio link. Antennas on the Wailuku exchange building are shown in Fig. 7.

Steps 2 and 3 of the expansion program are indicated in Fig. 8. Step 2 consisted of replacing the previous two overwater circuits to Kauai with new equipment for two circuits operating in the 76- to 108-mc frequency band. The previous circuit terminating at Manawahua, Oahu, was eliminated, and equipment for both of the overwater paths was placed at Mauna Kapu. A 900-mc repeater section for the second Kauai system was also installed as part of step 2. The two circuits to Kauai are not equipped with stand-by radio equipment. However, a transfer arrangement is provided so that the most important traffic will be switched to the operating system in the event of failure of one circuit. One of the Mauna Kapu buildings is shown in Fig. 9. A new station was constructed on the top of Kalepa ridge on Kauai for the equipment for the new circuits. This replaced an old station similar to that shown in Fig. 4, which was on the side of the ridge. The new station is about 250 feet higher than the old and provides a 13-decibel transmission improvement. Access to the Kalepa station is obtained via a sugar plantation jeep trail.

In the third step of the program, the two previously used systems to Maui were replaced by a single system. It provides 12 3-kc channels and utilizes the 900-mc repeater section previously installed between Ulupalakua to Wailuku. The 900-mc equipment installed during step 1 as a repeater section for the second Maui system was arranged to provide stand-by for the single remaining system.

In the fourth and final step of the program, the 100-mc equipment originally placed on the Oahu-Hawaii circuit was transferred to the circuits from Honolulu to Molokai and Lanai. New equipment with a wider modulation bandwidth was installed between Oahu and Hawaii. New circuits installed in step 4 are shown in Fig. 10 which also shows the interisland network upon completion of the entire expansion program. The circuit between Maui and Hawaii was removed shortly after the first step of the expansion program was completed.

To obtain the larger channel groups required for the new interisland network, much of the previously used channelizing equipment was retired, and new carrier systems were installed. As shown in Fig. 10, the type *42-6* equipment was retained and installed on the Kauai, Molokai, and Lanai circuits. In addition to the new carrier systems previously described for the Hawaii system, a new

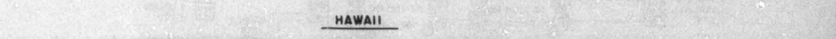

HAWAII
HONOLULU
TANTALUS
ULUPALAKUA
WAIKII
HILO
15 TFC
1 LEASED
2 TTY
42C
32E
76-108 MC
76-108 MC
42CR
9 TFC
1 LEASED
2 TTY
32E
6 TFC
WAIMEA
MAUI
HONOLULU
TANTALUS
ULUPALAKUA
WAILUKU
10 TFC
1 LEASED
1 TTY
456X
76-108 MC
920 MC
456X
10 TFC
1 LEASED
1 TTY
STANDBY
NEW EQUIP.
OLD EQUIP. RETAINED
KAUAI
HONOLULU
PENTHOUSE
MAUNA KAPU
KALEPA
LIHUE
8 TFC
1 LEASED
1 DIAL
2 TTY
4 TFC
1 DIAL
1 TTY
4 TFC
1 LEASED
1 TTY
42-6
900 MC
76-108 MC
900 MC
76-108 MC
42-6
4 TFC
1 DIAL
1 TTY
4 TFC
1 LEASED
1 TTY
8 TFC
1 LEASED
1 DIAL
2 TTY
MOLOKAI
HONOLULU
TANTALUS
KUALAPUU
KAUNAKAKAI
4 TFC
1 DIAL
1 TTY
42-6
76-108 MC
42-6
4 TFC
1 DIAL
1 TTY
LANAI
HONOLULU
TANTALUS
LANAI STA
LANAI EXCH
4 TFC
1 DIAL
1 TTY
42-6
76-108 MC
42-6
4 TFC
1 DIAL
1 TTY
T
R

Fig. 10. Interisland radio network upon completion of the expansion program

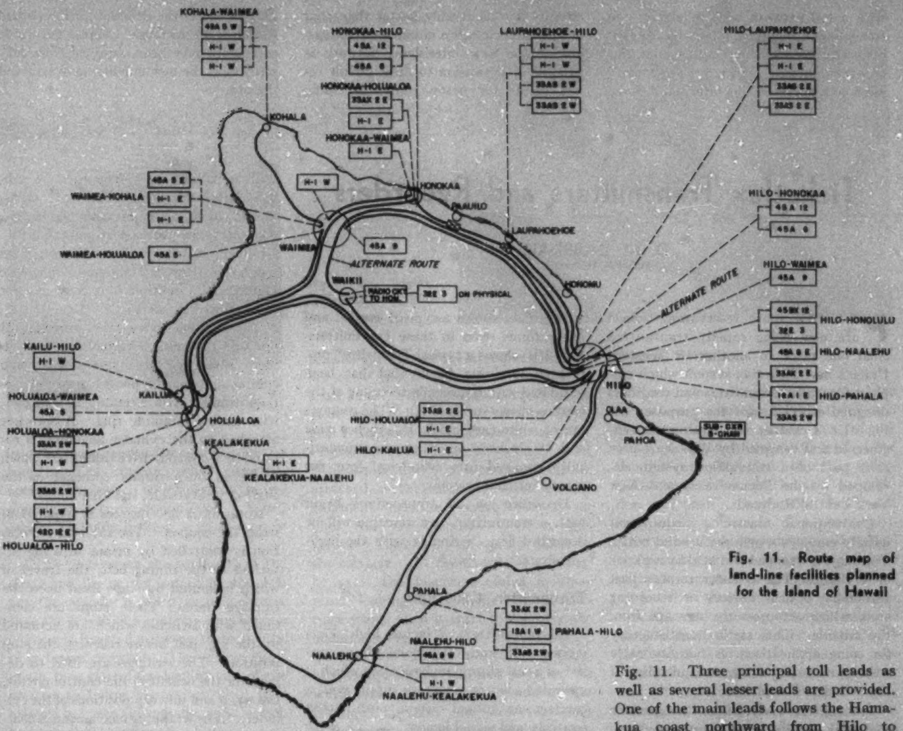

KOHALA-WAIMEA
45A 5 W
H-1 W
H-1 W
HONOKAA-HILO
45A 12
45A 6
HONOKAA-HOLUALOA
33AX 2 E
H-1 E
HONOKAA-WAIMEA
H-1 E
LAUPAHOEHOE-HILO
H-1 W
H-1 W
33AS 2 W
33AS 2 W
HILO-LAUPAHOEHOE
H-1 E
H-1 E
33AS 2 E
33AS 2 E
KOHALA
HONOKAA
PAAUILO
LAUPAHOEHOE
WAIMEA-KOHALA
45A 5 E
H-1 E
H-1 E
WAIMEA-HOLUALOA
45A 5
H-1 W
WAIMEA
45A 9
ALTERNATE ROUTE
WAIKII
RADIO CKT TO HON.
32E 3
ON PHYSICAL
HONOMU
HILO-HONOKAA
45A 12
45A 6
HILO-WAIMEA
45A 9
ALTERNATE ROUTE
45BX 12
32E 3
HILO-HONOLULU
45A 6 E
HILO-NAALEHU
33AX 2 E
15A 1 E
HILO-PAHALA
33AS 2 W
KAILU-HILO
H-1 W
KAILUA
HOLUALOA-WAIMEA
45A 5
HOLUALOA
HOLUALOA-HONOKAA
33AX 2 W
H-1 W
33AS 2 W
H-1 W
45C 12 E
HOLUALOA-HILO
HILO-HOLUALOA
33AS 2 E
H-1 E
HILO-KAILUA
H-1 E
HILO
OLAA
PAHOA
SUB-EXR S-CHAN
KEALAKEKUA
H-1 E
KEALAKEKUA-NAALEHU
VOLCANO
PAHALA
33AX 2 W
15A 1 W
PAHALA-HILO
33AS 2 W
NAALEHU-HILO
45A 6 W
NAALEHU
H-1 W
NAALEHU-KEALAKEKUA

Fig. 11. Route map of land-line facilities planned for the Island of Hawaii

Fig. 11. Three principal toll leads as well as several lesser leads are provided. One of the main leads follows the Hamakua coast northward from Hilo to Honokaa, one runs southward to Pahala, and the third follows the saddle route past the radio site at Waikii to the town of Holualoa on the west coast of the island. The saddle lead follows a path located between the two highest mountains of the island, Mauna Loa and Mauna Kea. Much of this route is over hard volcanic rock left as a residue from past volcanic activity on the island. Plans have been made to provide alternate routes for the circuits normally transmitted over the saddle lead since future volcanic activity of Mauna Loa may cause prolonged outages. When the circuits required for complete islandwide extended area service on Hawaii are completed, operators in Honolulu will be able to dial directly to all subscribers of the exchanges shown in Fig. 8.

l, 74
1
55

Lenkurt type *45BX* carrier system was installed on the Maui circuit.

Additional equipment was also obtained to provide for remote control and supervision of unattended radio sites from the Alakea exchange building in Honolulu. At the present time, personnel in Honolulu have full control over the systems operating to Hawaii and Maui. Work is now in progress on installation of similar control and supervision systems on the circuits to Molokai, and Lanai. When this work is completed, all of the radio sites will operate on an unattended basis. Since all radio terminals, with the exception of those on the Kauai circuits, are fully stand-by equipped, the entire radio network will be operated completely from one central location.

Integration With Other Toll Facilities

In addition to the interisland radio circuits described herein, the Hawaiian Telephone Company operates land-line toll facilities on all of the six islands. In all cases the intraisland facilities are fully integrated with the interisland network in such a manner that subscribers on any of the islands can make interisland calls quickly and easily.

The system now in operation provides facilities which make it possible for the Honolulu operator to dial all subscribers at exchanges on each island which are included in the dialing area for the exchanges at which the interisland circuits terminate. Also, additional carrier and land-line facilities are presently being installed as required to make each island a toll-free extended service area. When this is done, it will be possible for the Honolulu operator to dial any subscriber on any island and for any outer island operator to dial any subscriber on the island of Oahu.

As a typical example of the facilities existing on the various islands, a route map of the toll lines on Hawaii is shown in

Conclusions

Completion of the modernization and expansion program outlined in this paper has provided the Territory of Hawaii with a toll network with reliability and transmission quality comparable to that of any toll network covering a similar area on the mainland. In many cases, transmission quality of the interisland radio circuits is considerably better than that of the interexchange circuits on each island. The new interisland network is capable of expansion to meet circuit requirements for many years, as traffic growth causes the demand for interisland channels to increase. Many favorable comments have been obtained from subscribers as the new circuits have replaced the old.

Ticketfax Transmitters and Recorders

D. M. ZABRISKIE
NONMEMBER AIEE

PULLMAN space reservation tickets are now being rapidly transmitted from one place to another by Western Union's new Ticketfax system which employs facsimile transmitters and recorders designed especially for the purpose; see Fig. 1. Ticketfax is equipment engineered and designed by Western Union as a part of a reservation system developed by the Pennsylvania and New York Central Railroads.

Conventional facsimile equipment usually employs a cylinder around which a message is wrapped and held by a garter or a clamp. The cylinder rotates past stationary optical devices or recording styli, which act upon the message from the outside. This method, satisfactory for some applications, is not favorable where automatic loading and unloading of the cylinder is involved.

Transmitters and recorders designed for Ticketfax service employ internal scanning from within the cylinder. The message is automatically wrapped on the cylinder and scanned on the inside by rotating optical or stylus components.

Each machine is mounted upon a console, $29^1/_2$ inches high, which occupies a floor space of 18 by 18 inches. Each console, having two shelves, houses the control equipment associated with the machine it supports; the machines are slip-connected to the control circuits, permitting rapid removal and replacement.

Transmission facilities available for the initial installation prompted the selection of two speeds of operation, viz., 1,800 and 360 revolutions or scans per minute. Line feed for both speeds, based upon the use of a recording stylus of 0.008 inch in diameter, was established at 120 lines per inch.

Thus the high-speed sets, scanning at the line feed rate of $^1/_4$ inch per second, require 6 seconds to scan the $1^1/_2$-inch width of the coupon; the lower speed sets require 30 seconds. Loading and phasing add 2 seconds in each case.

Fig. 2(A) shows a typical coupon and other forms used in these transmitters. Fig. 2(B) shows a typical Teledeltos* preprinted receiving blank that has been validated with a facsimile recordng of the coupon shown in Fig. 2(A). The coupons are of white card stock, $1^1/_2$ inches wide and $4^1/_2$ inches long; they are automatically wrapped into cylindrical form for rotary internal scanning.

Inasmuch as all transactions begin with a transmitter, this machine will be described first, beginning with the high-speed unit; see Fig. 3.

Transmitter, 1,800 Rpm

All transmitters, 7 inches high, $17^1/_2$ inches wide, and $17^1/_4$ inches deep, consist of a cast aluminum base upon which is mounted a motor, a voltage supply unit, a carriage, an optical system, push-button controls, and signal lamps.

MOTOR

The shaft of a 1/75-horsepower synchronous motor protrudes at both ends. One end supports and drives a scanning mirror; the other end, equipped with reduction gearing, drives a chain in one direction to effect a line feed for the carriage. Also mounted on this end of the shaft is the rotary element of a phasing pulse generator.

VOLTAGE SUPPLY UNIT

The voltage supply unit powers the photomultiplier and is mounted on the transmitter base to utilize available space. Its proximity to the photomultiplier permits interconnection with short leads.

CARRIAGE

A carriage subassembly travels for line-feeding purposes from right to left, on linear ball bearings over a stationary shaft and is guided by roller bearings, engaged with a parallel rail. The carriage subassembly comprises a rotatable transparent lucite cylinder over which the coupon to be scanned is automatically wrapped, a wrapping mechanism with control switches, loading chute, magnetic clutch, and a multiple plug connector for internal wiring; see Fig. 4.

The transparent cylinder, threaded at one end to facilitate convenient replacement, is screwed into a steel sleeve which rotates in a bearing pressed into the carriage frame. The assembly is held endwise with a retaining ring. A sprocket attached to the cylinder assembly is rotated by a positive-drive timing belt which is driven by a similar sprocket on the shaft of a reversible, stallable motor.

Rotation of 380 degrees is required to wrap the coupon. The interval of rotation is controlled by means of a lug attached to the timing belt, the travel of which is limited by stops fixed upon the carriage frame. These stops are combined with switches which are actuated by the lug, just before reaching the stop position. The switches are used to detect, for the benefit of the control circuit, the wrap and unwrap positions of the cylinder. The wrapping mechanism is comprised of the transparent cylinder, a plastic web, a rotatable, spring-driven, retractile drum, and a loading chute.

For mounting purposes, the plastic web has one end folded over and stitched to form an eye. The eye is inserted into a slot cut in the transparent cylinder, and is retained by a length of straight wire inserted in the eye. The opposite end is clamped to the retractile drum, which is driven by an adjustable torsion spring and rotates in a direction tending to keep the web taut at all times.

The loading chute is a stainless-steel channel slightly wider than a coupon, open at the front end, adjacent to the cylinder, but closed at the back. A sensitive switch is mounted underneath the chute, the actuator of which protrudes

---

* Registered trade-mark, Western Union Telegraph Company.

Paper 55-503, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 29, 1955; made available for printing May 12, 1955.

D. M. ZABRISKIE is with the Western Union Telegraph Company, New York, N. Y.

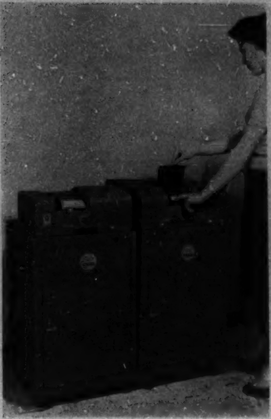


Fig. 1 (left). Ticketfax recorder and transmitter

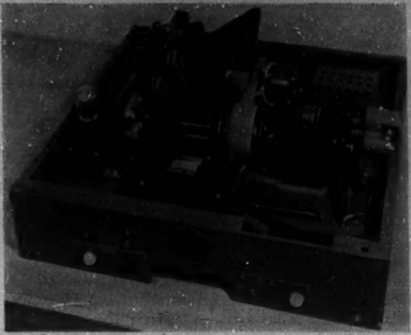


Fig. 3 (right). Transmitter 1,800 rpm

into the space eventually occupied by the coupon. The coupon, when inserted, operates the switch which permits the beginning of a transmission cycle. A call cannot be made if there is no coupon in the chute.

Distance between the back stop of the chute and the vertex formed by the web and the cylinder is 3/4 of an inch shorter than the length of a coupon; thus when a coupon is inserted between these points, it is bowed upward. The spring effect of the bow causes the leading end of the coupon to follow the vertex until it is gripped between the cylinder and the web and completely wrapped.

Carriage Line Feed

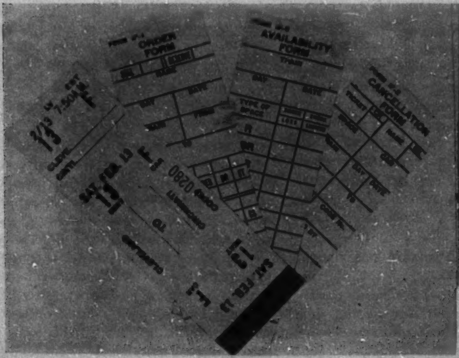
The carriage assembly is held against a bumper, in stand-by position, by an extension spring. A chain running on two sprockets moves parallel to the mirror shaft at a rate of 0.008 inch per revolution of the mirror shaft. It is driven by worm gear reductions taken off the motor shaft extension. A third sprocket, compounded with the armature of a magnetic clutch attached to the carriage, is held in engagement with the passing chain by an idler roll. In stand-by position this sprocket, rotated by the chain, idles on the carriage. At the instant of phasing, however, the magnetic clutch is energized, locking the sprocket to the carriage which then is carried along by the chain; see Fig. 5.

After the carriage has traveled laterally to the extent required to scan the width of the coupon, an actuator on the carriage operates an end-of-message switch which, through control circuit connections, de-energizes the magnetic clutch permitting the sprocket to idle backwards and the carriage to be retracted to its starting position by the extension spring.

Optical System

The purpose of the optical system is to focus a spot of intense light of 0.008-inch diameter upon the inside face of the wrapped coupon through the wall of the transparent cylinder and to collect reflections therefrom, again through the cylinder, and project those reflections to the light-sensitive element of a photomultiplier. On these transmitters all optical components are stationary on the base with the exception of one mirror which is mounted on the motor shaft; see Fig. 6.

A 32-candle-power tungsten filament lamp is used as an exciter. Light from this source is projected through a tube containing a condensing lens, an aperture, and an objective lens. The converging beam is reflected to the copy by two front surface mirrors, one rotating. The beam of light describes a circle on the wrapped coupon; this action in combination with the axial feed of the carriage scans the surface of the coupon in approximately 6 seconds.



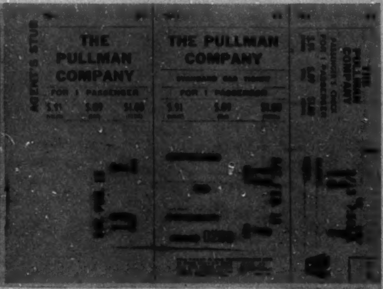
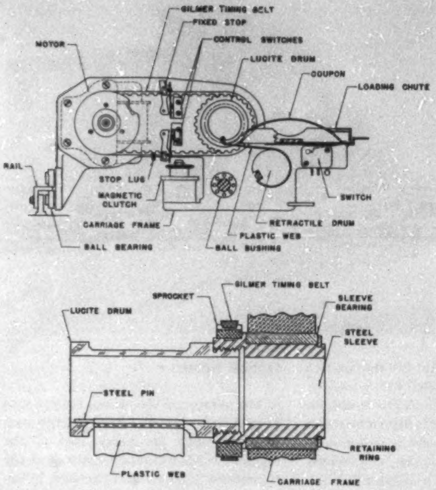


Fig. 2(A) (left). Coupon and miscellaneous sending forms

Fig. 2(B) (above). Recorded ticket


GILMER TIMING BELT
FIXED STOP
CONTROL SWITCHES
LUCITE DRUM
COUPON
LOADING CHUTE
MOTOR
RAIL
STOP LUG
MAGNETIC CLUTCH
CARRIAGE FRAME
BALL BEARING
SWITCH
RETRACTILE DRUM
PLASTIC WEB
BALL BUSHING
SPROCKET
GILMER TIMING BELT
SLEEVE BEARING
LUCITE DRUM
STEEL SLEEVE
STEEL PIN
RETAINING RING
PLASTIC WEB
CARRIAGE FRAME

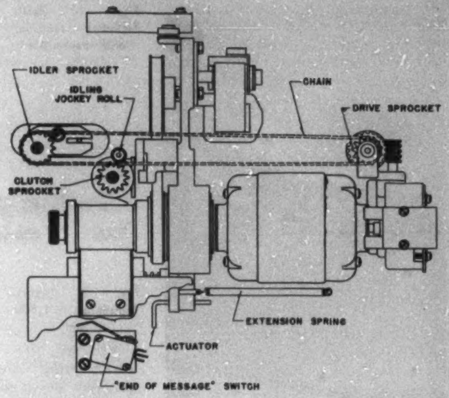

IDLER SPROCKET
IDLING JOCKEY ROLL
CHAIN
DRIVE SPROCKET
CLUTCH SPROCKET
EXTENSION SPRING
ACTUATOR
"END OF MESSAGE" SWITCH

Fig. 4 (left). Transmitter carriage

Fig. 5 (above). Carriage drive

A portion of the reflected light radiating through an aperture of the mirror shield impinges upon the rotating mirror from which it is reflected in a diverging cone, now limited by the concentric aperture comprising the end of the mirror shield, to the surface of a double Fresnel lens which collects the light and converges it onto the cathode of the photomultiplier. The rotary mirror is placed at an angle to prevent glare reflections of glossy tickets and of the inside wall surface of the cylinder from returning to the mirror; instead they are directed harmlessly in the opposite direction.

The exciter lamp, rotary mirror, and photomultiplier are separately shielded to prevent interference from internal transient light. Complete protection from exterior transient light is obtained with the main cover which is equipped with a hinged access lid for loading and unloading purposes. The lid is automatically latched so that it cannot be lifted during transmission.

Phasing-Pulse Generator

The phasing-pulse generator comprises a conventional magnetic pickup coil wound on a core having an air gap traversed by a permanent magnet which generates a pulse used for phasing and blanking; see Fig. 7. For transmitters and recorders to co-operate for phasing

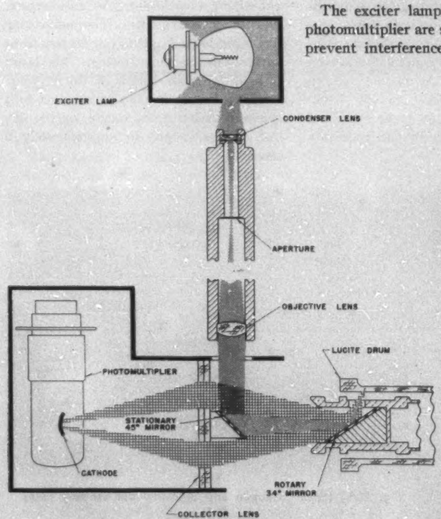

EXCITER LAMP
CONDENSER LENS
APERTURE
OBJECTIVE LENS
LUCITE DRUM
PHOTOMULTIPLIER
STATIONARY 45° MIRROR
CATHODE
ROTARY 34° MIRROR
COLLECTOR LENS

Fig. 6. Optical system

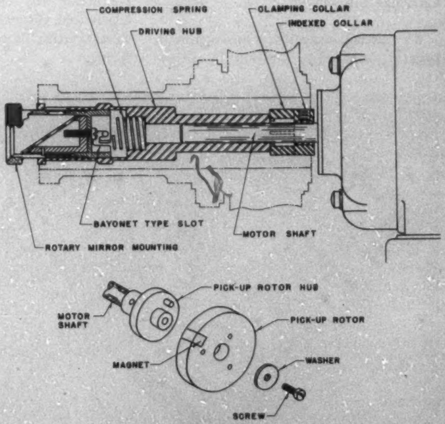

COMPRESSION SPRING
CLAMPING COLLAR
INDEXED COLLAR
DRIVING HUB
BAYONET TYPE SLOT
MOTOR SHAFT
ROTARY MIRROR MOUNTING
PICK-UP ROTOR HUB
MOTOR SHAFT
PICK-UP ROTOR
MAGNET
WASHER
SCREW

Fig. 7. Phasing rotor and rotary mirror

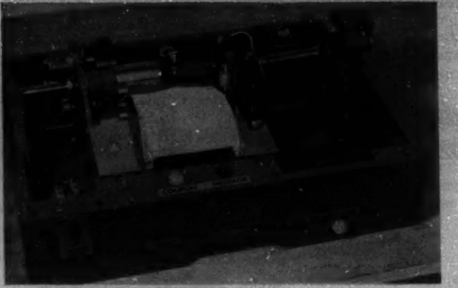


Fig. 8 (left). Recorder 1,800 rpm

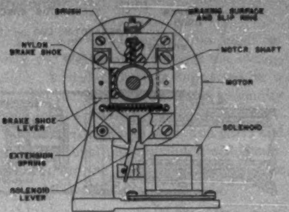


Fig. 9 (right). Stylus retractor brake

purposes, the magnet, imbeded in a brass disk, must be accurately located in relation to a motor pole in a specific position common on all machines. This relationship is established at the factory before the components are assembled.

Furthermore, if a remote transmitter should be powered from a phase of alternating current unlike that used to power an associated recorder, the machines will not co-operate in phasing, because of the time lag between the unlike phases. Therefore the pulse generator magnet disk is designed to be adjustable to three positions, each 120 degrees apart. Inasmuch as this is a field adjustment which must be made without disturbing the motor pole relationship established at the factory, the magnet assembly is made in two parts. The assembly comprises a steel hub with a key pin, and a brass disk containing the permanent magnet and having three holes, 120 degrees apart, matching the key in the steel hub. The steel hub, being pinned to the motor shaft at the factory, retains its relationship to the motor poles, while the brass disk can be adjusted to match any of the three phases of alternating current. At this point it should be noted that the rotary mirror, to maintain proper horizontal framing, must also be shifted to match the shift of the magnet.

Rotary Mirror Subassembly

Mountings for the rotary mirror have been designed to facilitate angular and lateral adjustment. For maintenance purposes, the mirror component may easily be removed without disturbing previously established adjustments. The mirror of this subassembly is held in place by a compression spring and a bayonet-type slot, requiring pressure and a slight twist to detach it from the driving hub.

The driving hub fits snugly over the motor shaft; the left end is slotted and a clamping collar is placed over it and keyed to the driving hub so that the hub and collar rotate together. A setscrew in the collar squeezes the slotted portion and securely grips the motor shaft. The clamping collar has a single engraved index line. Adjacent to the clamping collar, a second collar is fastened to, but is adjustable on, the motor shaft. It is engraved with three index lines, spaced 120 degrees apart. These indexed collars are used to gauge the shifting of the mirror to coincide with shifts made in the pulse generator, as previously described.

After the mirror mounting has been adjusted angularly to frame the ticket horizontally, it is fastened to the shaft with the clamping collar. The second collar is then moved so that one of its index lines coincides with the line on the clamping collar and is then fastened. The mirror mounting may now be moved in steps of 120 degrees to coincide with similar shifts of the pulse magnet without losing the framing relationship.

Transmitter, 360 Rpm

The slow-speed transmitter is identical with the high-speed unit with one exception. A shaft and bracket are mounted upon the cast pad normally used for the motor in the high-speed unit. A motor with a single shaft extension is mounted at right angles to the mirror shaft upon a second base pad provided for the purpose. Reduction spiral gears now couple the motor to the mirror shaft and, in as much as the line feed drive is obtained with gearing from the mirror shaft, the rate of feed is proportional.

Signals and Manual Controls

Each transmitter is equipped with four push buttons and three signal lamps. Two push buttons, each associated with a lamp of distinctive color, are used for starting. One, labeled coupon, operates similar signals on the called recorder, notifying the operator that a coupon is to be transmitted; the other, labeled message, serves notice that a message such as a space request form or handwritten message is ready for transmission.

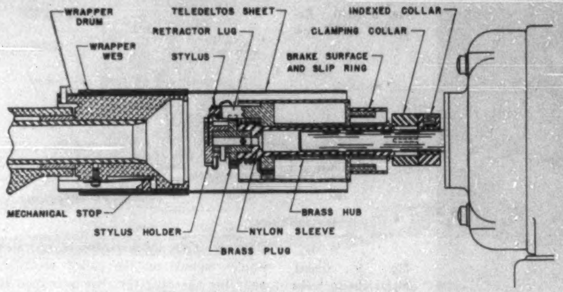
After a transmission is complete a lamp signal, labeled end-of-message, lights and must be extinguished by pressing the third push button which restores the transmitter to stand-by. The fourth push button, located obscurely under a hand hole in the base, is used as a manual end-of-message switch. This control, however, is used only when the operator wishes to halt transmission upon discovery of an error in the selection of matter being transmitted.

Recorders

Recorders, like the transmitters, are designed for two scanning speeds, viz., 1,800 and 360 revolutions or scans per minute. Recorders are 7 inches high, 17$^1/_4$ inches wide, and 11$^3/_8$ inches deep; see Fig. 8. The operation and mechanical arrangement of the recorder is similar to that of the transmitter. Many of its parts, such as motors, line feed drive, phasing-pulse generator, etc., are interchangeable. A recording stylus is used in place of an optical pickup system. A mechanical phasing clutch is used on the low-speed recorder in place of a phasing-pulse generator.

The ticket is a preprinted Teledeltos sheet 4 by 5$^1/_8$ inches on which an area in the lower right-hand section is reserved for recording. It is perforated along lines so that it is separable into three slips. The ticket is inserted in a loading chute; the left side of the leading edge abuts a mechanical stop and establishes phase relationship with the stylus. This stop is flexibly mounted inside the wrapper drum so that beyond 360 degrees of rotation the ticket pushes the protruding stop flush with the drum and the cylindrical form is not distorted.

The trailing edge of the ticket abuts a back stop in the chute. The space be-


WRAPPER DRUM
WRAPPER WEB
TELEDELTOS SHEET
RETRACTOR LUG
STYLUS
INDEXED COLLAR
CLAMPING COLLAR
BRAKE SURFACE AND SLIP RING
MECHANICAL STOP
STYLUS HOLDER
BRASS PLUG
NYLON SLEEVE
BRASS HUB

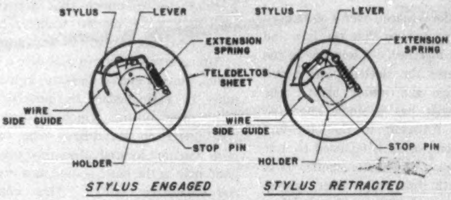

STYLUS
LEVER
EXTENSION SPRING
TELEDELTOS SHEET
WIRE SIDE GUIDE
STOP PIN
HOLDER
STYLUS ENGAGED
STYLUS
LEVER
EXTENSION SPRING
WIRE SIDE GUIDE
STOP PIN
HOLDER
STYLUS RETRACTED

Fig. 10. Stylus mechanism

tween front and back stop, being shorter than the ticket length, causes the ticket to bow upward. Spring effect of the bowed ticket is used to urge the leading edge forward until it is firmly gripped between the drum and the plastic wrapper web.

When the start button is pressed, a reversible, stallable motor drives the wrapping drum through a spur gear train, carrying the wrapper and the ticket along with it. As it is wrapped on the drum, the web is unreeled from a take-up drum under the loading chute.

The extent of rotation is limited by stops rotating with the gear train which engage a fixed stop on the frame. At this point the wrapper motor goes into a stalled condition, holding the ticket in the wrapped position for scanning. Upon unwrapping, the validated ticket is discharged into the loading chute.

Associated with the gear train are two cam-operated switches which are used to detect the wrap and unwrap positions for the control circuit. A detector switch with its actuator placed in the path of the leading edge of the ticket as it is being wrapped is used to inform the controls of the presence of paper for recording. If inadvertently the start button is operated causing the mechanism to wrap with no paper in it, failure of the switch to operate causes the drum to unwrap immediately upon operation of the wrap switch, thus restoring the mechanism to stand-by.

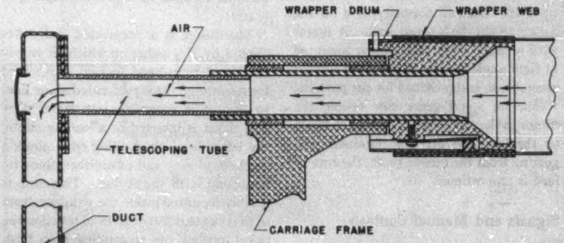

AIR
WRAPPER DRUM
WRAPPER WEB
TELESCOPING TUBE
CARRIAGE FRAME
DUCT
BASE

Fig. 11. Recorder exhaust

The wrapper drum supports, at the left side, only 1$^{3}/_{4}$ inches of the 4 inches of ticket width, leaving a substantial portion of the cylindrical tube unsupported and available for internal scanning. Rigidity, sufficient to support stylus pressure without distortion, is derived from the overlapping cylindrical form and the supporting effect of the drum and wrapper.

Stylus Assembly

Problems encountered in the design of a stylus mechanism to operate within the limits of a cylinder approximately 1$^{5}/_{16}$ inches in diameter, involved provision for such features as:

1. Light pressure.
2. Maximum stylus life.
3. Adequate insulation.
4. Retraction and release.
5. Rapid, convenient replacement of worn styli.
6. Maximum venting for the passage of air for exhaust purposes.

Vol. 7
Pt. 1
1955

The protruding stylus, if left in its normal recording position, would be damaged from interference with the paper during the wrapping and unwrapping operations. Therefore it is necessary to retract the stylus while in motion from the circle of contact with the paper. This is accomplished with a brake mechanism associated with the stylus assembly.

The stylus assembly comprises a brass hub mounted, and rotatable, on a nylon sleeve which serves the double purpose of an insulator and a dry bearing. The sleeve is eventually clamped to the motor shaft; see Fig. 9. Integral with the hub, a stainless-steel cylinder serves the dual purpose of a braking surface and a slip ring. The slip ring engages a brush and conducts the incoming facsimile signals to the rotating stylus. Acting upon the braking surface of the cylinder are two nylon brake shoes, pivotally mounted on levers so as to be self-centering. Again the nylon serves as a dry bearing and also as an insulator.

The brake levers are drawn together with an extension spring which causes the brakes normally to exert braking force upon the hub. This force tends to rotate the hub backward, with respect to the motor shaft rotation, increasing the tension of a torsion spring. One end of the spring is fastened to the hub, the other to the sleeve, thus indirectly to the motor shaft. The spring also serves to conduct the facsimile signals from the hub to the stylus. Brakes are released by the movement of a lever, actuated by a solenoid, which spreads the brake shoes apart; see Fig. 10.

A brass plug, fixed by pinning to the end of the insulating sleeve, serves as a mounting for a stylus holder. A portion of this plug is flanged and notched. The sides of the notch act as stops for a retractor lug protruding from the brass hub. In this manner the rotational effect upon the brass hub obtained by the alternate action of the brake shoes and torsion spring is limited.

The stylus holder is a detachable subassembly consisting of an L-shaped lever pivotally mounted and guided in a brass cap which fits over the plug on the end of the insulator and is snap-fastened thereto to facilitate easy removal. The lever rotates about a pivot and is biased toward a stop pin by means of an extension spring.

The stylus is formed of tungsten wire of 0.008-inch in diameter, eyeletted at one end and attached to the lever with a single screw. A wire extension from the lever supports the stylus sidewise. Pressure against the recording paper is derived entirely from the resilience of the tungsten form, and averages 5 grams. Vibration is prevented by utilizing the resilience of the tungsten wire to bear, in transverse direction, against the side guide.

Normally, the brakes are engaged and the stylus is held retracted while the cylinder is being formed, after which the brakes release, permitting the stylus to engage the recording paper. In laboratory tests, one stylus has recorded 20,000 coupons.

EXHAUST

Recording develops a certain amount of smoke and soot, which on these recorders is exhausted through a filter by a blower located in the console; see Fig. 11. The intake of the blower is coupled to the recorder by means of a duct and a telescoping tube inserted in the tubular wrapper.

This passage, extended further by the cylindrical form of the ticket, causes air to be drawn past the recording stylus with a considerable velocity to convey the smoke and soot to the filter. Other functions of the recorder such as line feeding, phase-pulse generation, etc., are identical with those of the transmitter, permitting use of identical parts.

Recorder, 360 Rpm

The slow-speed recorder is identical with the high-speed unit with one exception. A stylus shaft is supported in a bracket occupying the space normally used for the motor of the high-speed unit. The bracket supports reduction gears and a conventional phasing clutch which substitutes for the pulse generator used in the other machines.

SIGNALS AND MANUAL CONTROLS

An incoming call sounds a buzzer and lights one of two lamps, signifying the kind of message to be recorded. After the chute is loaded with the proper form, a push button is depressed to start recording. At the end of recording, the buzzer again sounds as an end-of-message lamp lights. Depression of a push button silences the buzzer, extinguishes the lamp, and restores the recorder to stand-by.

If a particular type of message can be anticipated, the operator may load the chute and depress an automatic answer relay after which the recorder is under the control of the transmitting operator. No further attention is required of the receiving operator until the buzzer sounds, indicating completion of the recording.

l. 74
1
55

Full-Wave Reversible-Polarity Half-Cycle-Response Magnetic Amplifiers

C. B. HOUSE
ASSOCIATE MEMBER AIEE

MAGNETIC amplifiers for servo control applications are becoming widely used and seemingly have a greater future ahead because of their inherent reliability and promise of small weight and volume.

The forerunner of the servo amplifiers was the simple reactor bridge wherein four reactor elements were connected across the supply in bridge form. When two opposite arms of the bridge saturated, a given polarity of alternating voltage was impressed across the load element. If the other set of arms had saturated instead, there would have been a reversal of the voltage polarity in the load element. Reference 1 states that the time delay in this system was somewhat longer than for a self-saturating circuit, and implies that the control power required is larger than the self-saturating circuit.

Later, self-saturating "doubler" circuits were connected across the two sides of a center-tapped transformer and the polarity of the alternating voltage applied to the load connected in the center leg determined by the circuit which first saturated.[1] This circuit had improved gain but did not posses half-cycle response. The desire for high response speed in servo applications later inspired a half-wave bridge-circuit magnetic amplifier with half-cycle response time, which used two cores per stage.[2]

Meanwhile, circuits were developed which at zero signal present unidirectional pulses at a rate of twice supply frequency to the controlled phase of a 2-phase motor.[3] The pulses are of equal average value so the fundamental supply frequency component is balanced out until a signal is presented to the amplifier. Upon introduction of a signal, chosen alternate pulses are caused to vary in the same direction. This produces a 60-cycle component in the output, which will cause motor rotation. Thus the motor sees a supply frequency component when a signal is presented and a d-c component at zero or small signal values for rotational damping. The unidirectional system has a high audible noise level at zero signal which is especially noticeable in a 60-cycle system.

The unidirectional output magnetic amplifier idea was combined with the half-cycle response concept at the Naval Research Laboratory and resulted in a unidirectional type of servo amplifier with half-cycle response time.[4]

The theory of half-cycle-response amplifiers has become widely circulated.[5] Suffice it for now to repeat that in this type of amplifier the firing angle of the saturable core is fixed uniquely during the previous half-cycle of applied carrier volt-

---

Paper 55-537, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 30, 1955; made available for printing May 13, 1955.

C. B. HOUSE is with the Naval Research Laboratory, Washington, D. C.

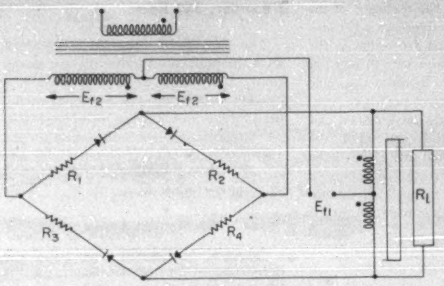


Fig. 1 (above). Basic ring modulator circuit

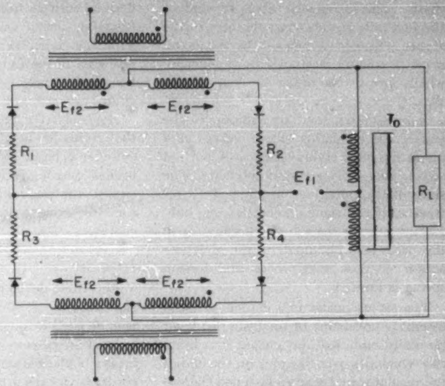


Fig. 2 (right). Ring modulator with series switching-voltage sources

age, irrespective of the history of voltages and/or currents applied to the core prior to the initiation of the control half-cycle. Thus the speed of response is limited only by the carrier frequency and not by circuit components.

Seemingly, there still remains room for a high-gain full-wave reversible-polarity half-cycle-response magnetic amplifier for servo applications, so several circuits are presented for the purpose of filling that gap. The consideration of the circuits should not, however, be limited to servo applications since they are inherently suppressed carrier modulators capable of a-c or d-c output with either a-c or d-c input signals, which should be useful in other fields.

Background

The concept of the present circuit as a modulator follows very easily since it was originally developed from ring modulator circuitry.[8] The basic ring modulator configuration, Fig. 1, has been characterized by the rectifier-resistor ring with external voltages for switching and modulating purposes. In the basic circuit form, the modulating low-frequency signal voltages E_{f1} cannot rise instantaneously higher than one-half the switching voltage E_{f2}, without unblocking one of the ring rectifiers which is supposedly nonconducting.

By using two center-tapped switching-voltage sources, inserted in series with the rectifier ring (Fig. 2) positive blocking action is achieved for all values of $E_{f1}<E_{f2}$. This blocking action will be seen to be of prime importance in adapting the ring modulator for high-speed magnetic-amplifier circuitry.

Geyger[7] utilized the flexibility of the ring modulator to devise push-pull a-c magnetic amplifiers. Output windings were inserted in series within the rectifier ring of Fig. 1. The modulating voltage E_{f1} appears within the ring as an unbalance voltage due to the output cores saturating at different times in the gating half-cycle. However, with the switching voltage still external to the ring, the improper unblocking of the rectifiers prevents full control of the resetting cores and prevents high-speed amplifier response.

Fig. 3 shows a magnetic amplifier (devised by D. G. Scorgie of the U. S. Naval Research Laboratory) based upon the ring modulator of Fig. 2. The square designation for the core of the saturable cores and output coupling transformer signifies that the core material is of the high-remanence type such as that in grain-oriented 50 per cent (%) nickel, 50% iron materials. The coupling transformer is constructed of high-remanence material to preclude reaction of the coupling transformer upon the saturable cores from stored energy feedback.

The object of this circuit is to decouple the resetting cores from the effects of the output pulses. Suppose, for instance, that cores 1 and 2 are on their gating half-cycle and cores 3 and 4 are resetting. Resistor R_c is chosen so that when an output pulse appears, half the output voltage appears across R_c. The polarities of the voltages across T_0 and R_c are such that the voltage from ground to the center tap of transformer T_2 remains exactly zero. Therefore, the resetting core windings are free to assume any value between zero and E_{f2}, the desired range. To prevent the output pulse from coupling to the resetting cores through the control circuit, a feedback voltage E_b is introduced from the output coupling transformer T_0. The completed circuit gives full response in one half-cycle of the supply frequency.

As the circuit stands in Fig. 3, cores 1 and 2 will be proceeding towards saturation simultaneously, so if both cores reach saturation during their gating half-cycle, series current-limiting resistors would be required in each arm of the bridge. This represents a power loss which adds nothing to the operation of the circuit. The simplest circumvention of this loss would be reversing the rectifiers in series with cores 2 and 3. This eliminates the short-circuit path between end terminals of the supply transformers, and shifts the burden of current-limiting duties to R_c, which already serves a useful purpose in the circuit.

After these changes, the circuit, Fig. 4, has become a functioning circuit wherein the direction and duration of current flow through the load resistance is controlled by magnetic amplifier cores, and interaction of the various cores has been eliminated. This elimination of interaction permits complete control of the resetting cores through its individual control winding.

Control Circuits

The connection of the control windings has previously been deliberately ignored since the development of their purposes is a separate consideration. Two methods of control will be developed, each of which has advantages and disadvantages. The user of the circuitry may decide on the preferred control method for his particular application. For purposes of differentiation, one will be called the series feed and the other the parallel feed.

Vol
Pt,
195

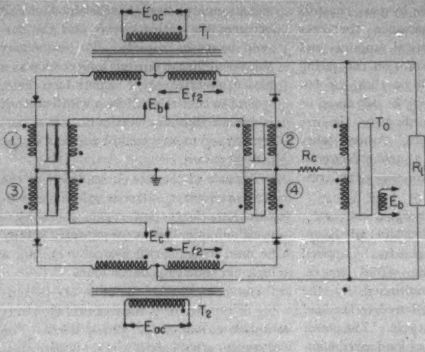


Fig. 3. Basic magnetic-amplifier controlled ring modulator circuit

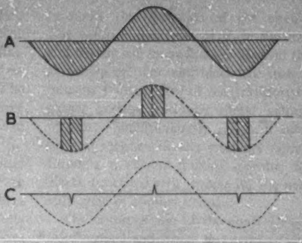


Fig. 5. Ouptut current wave forms of magnetic amplifier

A—At maximum signal voltage
B—At half-maximum signal voltage
C—At near minimum signal voltage

Slight differences will appear in the gating circuits but basically they will be similar.

A preliminary consideration in the development of the control circuitry was the desired shape of the output wave form. Measurements show that in 2-phase servomotor operation, the component of control-phase current which is effective in producing motor torque was that fundamental frequency component which is in a 90-degree angular relationship with the reference phase current. This limitation requires current pulses in the control phase to center about the peak voltage point rather than occurring near the end of the gating half-cycle. To obtain the desired phase relationship within the controlled phase for small output currents, the control circuit was designed under the assumption that at zero signal voltage, the gating cores would simultaneously saturate at 90 degrees. In the circuit shown in Fig. 4, for the polarities shown, cores 2 and 4 would be on their gating half-cycle. If the two cores saturate at the same instant, the currents through the two halves of the output coupling transformer, T_0, will be of equal magnitude and of opposing direction so that the net effective current for magnetizing the coupling transformer will be zero. Thus no voltage will appear across the load.

Upon introduction of a signal, the two resetting cores will not see the same resetting volt-seconds. On their next gating half-cycle, one core will advance its firing angle a certain number of degrees and the other core will have its firing angle retarded a like amount. During the period subsequent to one core's gating and prior to the other core's firing, a voltage will appear across one-half of the coupling transformer. Because of transformer action, twice this voltage will appear across the load. The polarity of this voltage pulse will be fixed by the core which fires first. The current flowing through the load for various output levels woud appear as shown in Fig. 5 if the signal had the same polarity for its different magnitudes.

SERIES-FEED CONTROL

The simplest way to introduce a signal into the amplifier control windings is to connect them all in series with the signal E_s with the use of the winding senses shown in Fig. 6. These senses are used for a-c control voltages. It is evident that when the turns ratios on all cores are identical, the voltages transformed from the gating windings will sum to zero until one core saturates. When any gating core saturates, its transformed voltage drops to zero and the remaining gating core would tend to act upon the control circuit's resetting cores. By coupling back a voltage from the output coupling transformer of the same polarity and magnitude as that which 'vanished

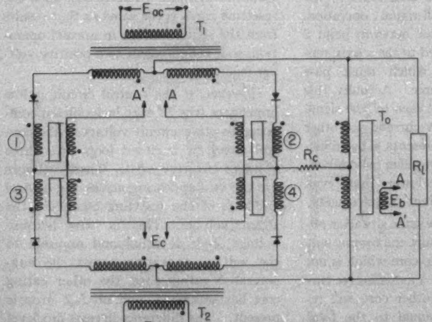


Fig. 4. Full-wave reversible-polarity magnetic amplifier

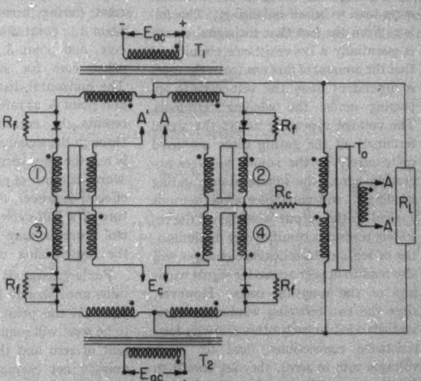


Fig. 6. Series-feed-controlled magnetic amplifier

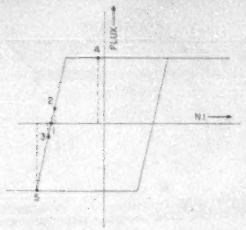


Fig. 7. Idealized hysteresis loop for toroid cores

through the terminals A-A', no transient disturbance will appear in the control circuits. The control circuits of the resetting cores will be influenced only by the signal voltage irrespective of the state of the gating cores.

The resistors R_f shown by-passing the load-circuit rectifiers, serve to set the no-signal flux level in the resetting cores at such a point as to produce firing at the 90-degree point in the gating half-cycle for the reason previously covered. They may do this since the direction of power supply voltage on the two nongating cores will tend to reset them if there is current leakage through the rectifiers. This leakage is increased to the proper value through the by-pass resistors.

The voltages appearing across the control-winding terminals of the resetting cores will be of opposing polarities, as were the gating-core voltages. Thus, in the absence of signal voltage, the net voltage sum across the signal windings will be zero and no current will flow through the signal source. It is interesting to note that in the no-signal condition the cores tend to be self-balancing. This follows from the fact that the signal circuit is essentially a low-resistance circuit and that the amount of flux change in the cores is dependent upon the volt-seconds appearing across the winding terminals. The voltage appearing across the signal terminals of the gating cores are fixed quite surely by the supply voltages appearing across the low-resistance gating windings. Any unbalance of volt-seconds applied to the output windings of the resetting cores as a result of core dissimilarities or rectifier back leakage changes will also tend to appear across the signal windings of the resetting cores. However, since the two resetting windings are essentially tied to each other through low-resistance connections, (and all other voltages sum to zero), they are forced to feel the same volt-seconds. This implies that even for dissimilar cores or for nominal changes in the by-pass resistors and rectifier back resistance, the cores will still be reset identical amounts and will fire at identical times on the gating cycle. The effect will be that the 90-degree firing position will be advanced or retarded slightly according to the net change in characteristics. Among other effects, this action is a strong factor in canceling output drift for zero signal voltages.

Upon application of a signal voltage E_c, the net difference in flux level of the resetting cores will be equal to the applied volt-seconds from the signal source. These volt-seconds multiplied by the turns ratio will be transferred to the load on the next gating half-cycle. The power gain will be a function of load circuit impedance and the dynamic impedance of the control circuits. This control-circuit impedance as defined by E over I will be comparatively high because of the circuit configuration.

At zero signal the magnetizing current for resetting the cores will not pass through the signal circuit, but will be supplied through the rectifiers and resetting resistors. For small values of signal voltage, the current flowing through the signal source will be the absolute difference between the magnetizing current required to reset the core slightly more than half way, and that required to reset the other slightly less than half way. Theoretically, for square hysteresis loop cores, this difference would be a function of the slope of the magnetomotive force curve between the two points on the hysteresis curve reached by the cores. See Fig. 7. The two hysteresis loops are assumed to be identical and superimposed. Each of the two cores would reset during zero signal operation to point 1. Point 2 is the final reset of one core, and point 3 the final reset of the other core for small-signal operation. The horizontal distance between point 2 and point 3, as referred to the x axis, represents the current which must pass through the signal source. Actually, this is not the only current seen by the signal source because of eddy-current and other effects; however, it presents a basic picture of a half-cycle amplifier whose control currents may be lower than even the magnetizing current requirements.

For larger signals the same situation obtains until shortly before maximum output. At this point the core which is not to be reset will require a net effective current of zero and the other core will require a net current equal to the total magnetizing current, or the difference between points 4 and 5. Of course, for signals greater than this the signal circuit becomes the gating circuit and the current increases rapidly. To counteract this possibility, a current limiter such as a biased rectifier or a square-loop core, wound so that the half-loop width current is greater than the maximum signal current requirements, could be placed in the control circuit.

Because of the low current flow in the control circuit, this device exhibits a comparatively large power gain for a half-cycle amplifier and concomitantly it may be used with control sources requiring a high dynamic impedance load.

One complicating factor in the designing of the circuit is the reaction after saturation of one of the gating cores. For instance, after core 2 in Fig. 6 saturates, the voltages of the two halves of the supply transformer are applied in series across the gating circuit of core 1. This is double the voltage applied prior to saturation and must be taken into account when fixing the size of the resistor by-passing the gating circuit rectifier for initial reset. The over-all effect of this action upon the circuit seems to be compensated for in the presence of a signal. If, for instance, a signal is present such as to fire core 2 earlier than the 90-degree angle, the double resetting voltage would be applied to core 1 earlier in the half-cycle, which would tend to reset the core more.

If there is a high-impedance element in series with the control circuit (high impedance being a vague phrase meaning only that any current flow which would tend to result from the double voltage would cause an appreciable voltage drop across the impedance), some added reset of core 1 would occur. During the same half-cycle the application of double voltage to core 3 from saturation of core 4 would be delayed. This action upon the resetting cores is the same as that results from the signal voltage in normal operation, so it is equivalent to a positive voltage feedback.

However, if the control circuit is low impedance (the IZ drop insignificant compared to other circuit voltages), the control circuit is a closed loop with three voltages effective in it. These voltages are the two appearing across the control windings of the resetting cores and the signal voltage. This is true because voltage A-A' is equal and opposite to the voltage appearing across the non-saturated gating core, the other gating core has saturated, and no I_mZ drop is present. The difference in reset flux level must be fixed as equal to the volt-seconds in the signal. The effect of the early double voltage appearing across one resetting core is to tend to shift the firing pulse into the later part of the cycle without changing the volt-second available to the load. This tendency to shift is minimized by the high-resistance values in the resetting circuit on the gating side. If one core saturates prior to the other and double voltage appears across the associated gating winding, this double voltage will appear in the control circuit and result in some current increase. Any slight change in current flow in the control circuit will be reflected into the gating winding as a large change in resistance drop, and consequent reduction in terminal voltage.

The control circuit heretofore discussed has been predicated upon a-c control signal and a-c output current. The circuit may be easily altered to give an alternating current of reversible phase in the load with reversible polarity direct current applied to the control circuit. This change could be simply made by reversing the windings on cores 1 and 3, or 2 and 4. At some sacrifice in efficiency and available output, the output coupling transformer may be replaced by coupling resistors and the compensating resistor R_c eliminated. This can give unidirectional output pulses with a polarity dependent upon signal polarity. Again the control signal may be either alternating or unidirectional voltage as fixed by the control-winding connections. Under these various conditions of operation, the feedback voltage to terminals *A-A'*, Fig. 6, would be resistance-connected or transformer-coupled depending upon the type of output voltage.

Parallel-Feed Control

To eliminate all reaction of the gating cores on the resetting cores, the parallel-feed circuit shown in Fig. 8 was evolved.

The gating circuits remain as in the series-feed amplifier except for the elimination of the by-pass resistors which established the initial core reset bias. The reset bias which will give a 90-degree firing angle on all cores in the absence of signal is now supplied by voltage E_r. With the connections shown, cores 1 and 3 will be resetting while cores 2 and 4 are gating, and E_r is adjusted to furnish exactly one-half the number of volt-seconds which were available to the resetting cores during their gating half-cycle. The magnetizing current for the resetting cores will flow through the secondary of the signal coupling transformer T_c but will have no effect upon the signal source due to the cancellation of magnetomotive forces.

If signal voltage E_c is impressed upon the primary of the signal coupling transformer, the transformed voltage will be of such a polarity as to aid E_r for one resetting core and subtract from E_r in the other core by equal volt-seconds.

Since the magnetizing current for the resetting cores is supplied by E_r at zero output levels, the signal source will only have to supply the differences in magnetizing current as previously explained and illustrated in Fig. 7. This factor in turn reflects a relatively high impedance into the control circuit.

Switching rectifiers r_{s1} and r_{s2} are included in series with E_r to decouple the gating cores from the signal transformer. Since E_r is a sinusoidal wave of one-half the volt-second content of the transformed gating voltage, the instantaneous values of the transformed voltage in the control circuits of the gating cores will be larger than E_r. This voltage would be of such a polarity as to cause current to flow through the rectifiers 1 and 2, or 3 and 4, to upset the operation of the circuit. The polarity and value of E_s in the control circuit will be such as to shut off the switching rectifier and block the instantaneous voltage sum of the transformed gating voltage and any signal voltage present. Simultaneously, E_s will open up the rectifier in the resetting circuit through a current-limiting resistor R_s. These functions will reverse on the next half-cycle of applied voltage.

The transposition of the control leads from cores 1 and 2 was necessary to give alternating output voltage through the output coupling transformer. Direct connection would cause cores 1 and 2 to saturate first during alternate half-cycles thus producing unidirectional current pulses through the corresponding half of the coupling transformer. This in turn would drive the core to saturation in one direction and consequently drop the transformer out of the circuit.

If reversible polarity unidirectional output current is desired, the coupling transformer may be replaced with a coupling resistor and the control circuit reconnected directly. If unidirectional voltage control is desired, the signal transformer may be replaced with a high-resistance coupling resistor and the signal voltage connected across the ends of the coupling resistor. This will entail a loss in efficiency in the control circuit since the coupling resistor will be in parallel with the cores across the signal source.

By applying the concept of winding terminal voltage control,[8] the parallel-feed magnetic amplifier may be altered for use with low-remanence narrow-hysteresis loop magnetic-core materials. See Fig. 9. The change has consisted of reversing all rectifiers in the control circuit, and connecting by-pass resistors across the gating circuit rectifiers to give full reset

74
1
55

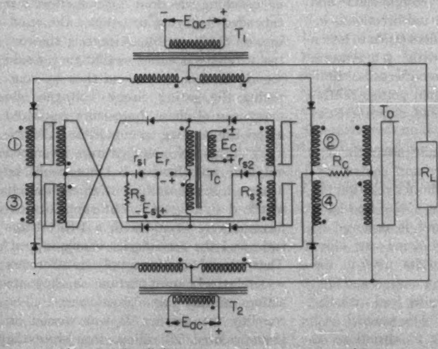


Fig. 8. Parallel-feed-controlled magnetic amplifier

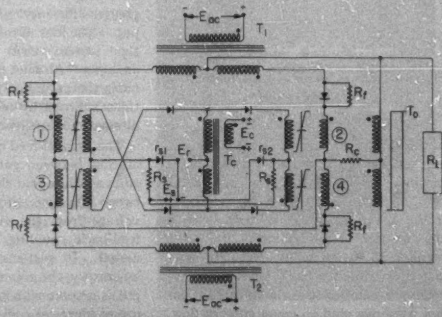


Fig. 9. Parallel-feed-controlled amplifier with low-remanence magnetic cores

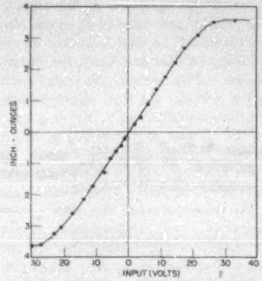


Fig. 10. Stall-torque versus 60-cycle component of signal voltage, input impedance > 40,000

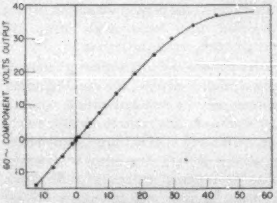


Fig. 11. Voltage transfer characteristics of amplifier with series-feed signal, load resistance 96 ohms and input impedance > 40,000

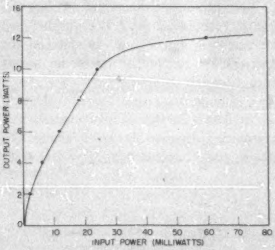


Fig. 12. Power transfer characteristics of series-feed amplifier

in the absence of bias and signal voltages. In no-signal operation, the voltage at the terminals of the control windings of the resetting cores (as set by leakage currents in the gating windings) will tend to be in the direction of conduction of the series rectifiers. The switching rectifiers associated with the resetting cores are in the conducting state so the voltage in the control windings tend to cause current flow. This current flow in turn causes the supply voltage to be absorbed as a current resistance drop in the by-pass resistors of the gating circuit. The net result is that no appreciable voltage appears across the winding terminals of the resetting cores (therefore no reset) when E_r and E_c are zero. Any voltage in series with the resetting winding which opposes the current flow will permit that same voltage to appear across the winding terminals, before current flow and consequent cancellation of flux change occurs. E_r in normal operation is fixed at the value of volt-seconds required to reset the core half way, as before. Any signal voltage applied through the signal transformer will then add to the reset in one core and subtract from the other.

This circuit as presented is entirely practical but no great advantage in gain is realized by the use of narrow-loop materials. This is partly because of the method of operation. Since the signal source has only to supply the difference in magnetizing currents, the power input is a function of the slope of the hysteresis loop and also a function of the change in loop width with frequency. The basic loop width is of only minor importance until just before full output. A secondary effect with narrow-loop cores, as a result of the lower maximum flux level, is that the number of turns has to be increased, which in turn increases the resistance in series with the load by approximately the square of the ratio of the maximum flux levels. This increase of internal resistance in the gating windings results in much lower power output levels. In one amplifier, substituting the same size supermalloy cores in place of square-loop cores resulted in the same total gains.

Comparison of Control Circuits

The series-feed control circuit has the advantages of fewer components and greater efficiency with unidirectional input. The lack of rectifiers tends to lessen the system drift with temperature changes. Absence of rectifiers also eliminates the minimum signal voltage limitations imposed by rectifier characteristics at small voltage levels.

The parallel-feed circuit makes the separation between gating and control circuits absolute and eliminates reaction between the two circuits. No feed-back voltage from the output is required so feed-back coupling problems are eliminated. In systems so far used in the laboratory, the gain for a-c input and output is greater with parallel feed with the use of identical cores. The parallel feed system is less sensitive to variations in core materials than the series feed. The use of low-remanence narrow-loop core materials has been more successful with parallel-feed control circuit than series feed circuits because of this lack of sensitivity. The limitations set on minimum signal voltage by the rectifier is modified by signal transformer coupling since the turns ratio may be varied to suit the signal.

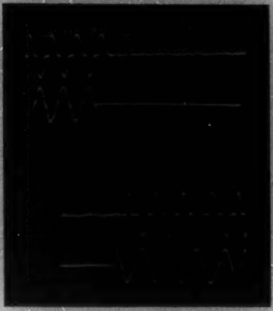


Fig. 13. Response time of series-feed magnetic amplifier

Top—Turn-off time. Lower trace is signal voltage, upper trace is output voltage
Bottom—Turn-on time. Trace order the same

Reduction to Practice

Series Control

Four grain-oriented 50% nickel, 50% iron saturable core toroids were wound so as to saturate with 60 volts at 60 cycles applied to the gating winding. The control winding on each core was wound to saturate at one-half this voltage. Only four cores of the desired size and material were immediately available so no selection or matching was done. Some effort was expended in trying to balance the windings on the cores by connecting the control windings in series so that they would see identical currents, and then by comparing the gating circuit voltages. So many out-of-phase harmonics appeared across the winding terminals that the primary frequency voltages were swamped out and no effective balance could be achieved, so the number of turns was left at the design figure. The supply transformers were wound with a 1-to-1 turns ratio and the secondaries center-tapped. The material used for supply transformers was relatively unimportant as they are acting simply as transformers. The coupling transformer T_s was wound on grain-oriented 5% silicon iron since the main requirement was high remanence.

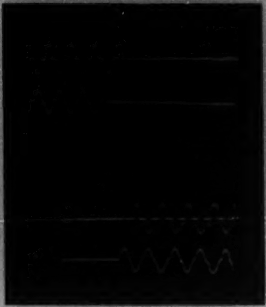


Fig. 14. Response time of series-feed magnetic amplifier with load resistance four times greater than optimum. See subtitle of Fig. 15

The cores were then connected in the circuit shown in Fig. 6 with the use of *1N151* germanium junction load rectifiers to get the required current capacity. The first four rectifiers available were placed at random in the bridge arms. However, the by-pass resistors were individually adjusted with the circuit in operation to give 90-degree reset of each core with open control circuit. Upon closing the control circuit, the self-balancing of the cores as previously described took place and further slight adjustments of the by-pass resistors were required.

The values of load resistance and compensating resistance for maximum power output were empirically determined with the aid of an electrodynamometer type of wattmeter. Transfer characteristics and response time measurements were then made for various circuit parameters.

Since one use visioned for this circuit was as a servo amplifier, one series of measurements was based upon the 60-cycle components of input and output voltages and currents as measured with a harmonic wave analyzer. This series included stall-torque measurements on a 2.5-watt 2-phase servomotor for given input signal voltages. The results of the stall-torque versus input voltage measurements are given in Fig. 10. The results are not necessarily optimum, since the amplifier was not designed for this motor, but they are indicative of the results possible with this circuit.

Fig. 11 shows the 60-cycle components of the output voltage when a resistance load of 96 ohms was placed across the output terminals. The input impedance, as calculated by dividing the 60-cycle component of signal circuit voltage by the 60-cycle component of signal circuit current, varied over a range of 20% for various outputs but was greater than 40,000 ohms at all outputs less than the maximum output voltage.

Simultaneously, with these measurements on the resistance load, readings were taken of rms input current as read on a thermal ammeter and indicated rms signal voltage as given by rectified average reading meter. Output power was read on an electrodynamometer type of wattmeter. Curve of the input power versus output power is given in Fig. 12. It is seen that power gain varied from 400 to 1,000 for outputs of less than 10 watts when measured by these methods.

Oscilloscope pictures were taken on a dual-beam scope to check the response time of the amplifier for both on-time after application of signal and off-time after release of signal, since amplifiers are often characterized by a fast attack time but a slow decay time. Results are shown in Fig. 13. The uppermost trace shows the output voltage after cessation of the signal voltage trace shown directly below. The bottom two traces are similar indications of turn-on time. The small pips remaining upon the output voltage wave with no signal are indicative of small core and component unbalance.

Theoretically, the relation of the value of the load resistance to the compensating resistance R_c should maintain a constant value, so concern was felt because of the effect of load impedance variation upon response time of the amplifier. A load resistor of four times the optimum value used for Fig. 13 was substituted in the output, and the oscillograph traces in Fig. 14 resulted. The order of traces remains the same as before: the topmost is the output voltage and immediately below is the signal voltage during turn-off. The bottom two traces are for turn-on time.

As a demonstration of an extreme case of unbalance, the compensating resistor R_c was short-circuited out and response time pictures taken. These traces are shown in Fig. 15. The order of traces remains as before. Observations of the circuit with loads less than optimum were also made and no degradation of response time was noted.

Parallel-Feed Circuit

The same cores used in series-feed tests were reconnected as shown in Fig. 8. The rectifiers used in the control circuit were *1N93*. Several other rectifiers would have been suitable but these were chosen because of availability and possibility of accidental high currents during experimentation. The tests for response time with various circuit parameters were

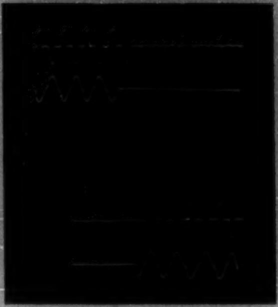


Fig. 15. Response time of series-feed magnetic amplifier with no compensating resistor.

Top—Turn-off time. Upper trace is output voltage, lower trace is signal voltage
Bottom—Turn-on time. Trace order the same

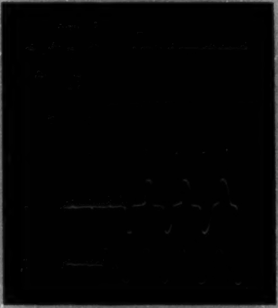


Fig. 16. Response time of parallel-feed magnetic amplifier.

See subtitle of Fig. 15

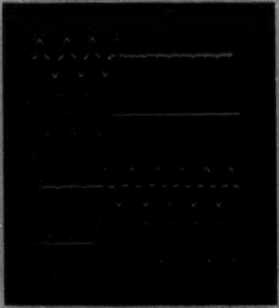


Fig. 17. Response time of parallel-feed magnetic amplifier without compensating resistor. See subtitle of Fig. 15

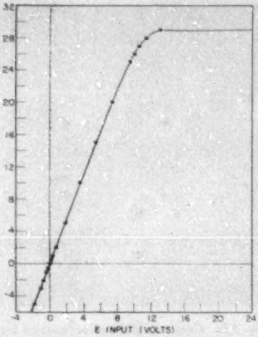


Fig. 18. Transfer characteristics of parallel-feed magnetic amplifier. Sixty-cycle components measured. Input impedance approximately 20,000 ohms, load resistance 75 ohms

repeated with the same results as before. Fig. 16 shows the turn-off and turn-on times of this circuit. The upper of each set of traces is the output voltage and the lower trace the signal voltage. Fig. 17 shows turn-off and turn-on time for the amplifier without the compensating resistor. Turn-off time is shown in the upper set. The transfer characteristic of the 60-cycle components of input and output voltages of the parallel-feed amplifier are shown in Fig. 18. This corresponds to the characteristic shown in Fig. 11 for the series-feed amplifier. These two curves indicate that the parallel-feed amplifier has from two to four times the power gain of the series-feed amplifier. Similar power gain differences also existed for rms quantities. In the foregoing tests, input impedance was set at approximately 20,000 ohms by selection of the input transformer turns ratio.

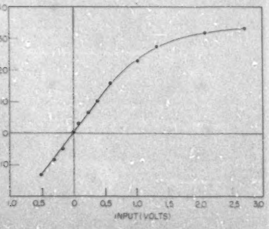


Fig. 19. Transfer characteristic of parallel-feed magnetic amplifier using supermalloy cores. Load resistance 360 ohms, input impedance approximately 1,000 ohms

In later tests the question of current limiting was examined for cases where signal levels became abnormally high. The input impedance, as defined by the ratio of the 60-cycle components of input voltage and current, was useful in determining loss of sensitivity where series-resistance limiting was employed. With series resistance added, the signal voltage divided approximately as though the control windings were linearly resistive, up to the point where the signal voltage saturated the amplifier. A current-limiting means which does not sacrifice so much gain in the operating region consists of the addition of a high-remanence toroidal reactor in series with the input, as previously discussed.

The results of using narrow hysteresis loop cores in these circuits were investigated next by substituting toroidal cores of 5% molybdenum and 79% nickel of the same physical size for the broad-loop cores. These cores were wound so as to saturate at the same impressed voltage as the previously described cores.

Because of the lower flux density, the number of turns was doubled and the wire size halved, in comparison with the grain-oriented nickel-iron cores. This decreased maximum output power to less than a quarter of that obtainable from the oriented nickel-iron core of the same dimensions. These molybdenum nickel cores were connected into the circuit of Fig. 9, with parallel signal feed. Sixty-cycle component transfer characteristics were obtained and are shown in Fig. 19. As previously stated, the improvement in magnetizing current requirements for these molybdenum nickel cores has no appreciable effect upon the gain in this circuit since the slope of the hysteresis loop is the controlling factor. That theory is reflected in the transfer characteristic curve of Fig. 19, which closely parallels the shape of the upper right-hand quadrant of the hysteresis loop of this material. The power gain is close to that obtainable with grain-oriented nickel-iron cores in parallel-feed circuits at lower outputs and decreases at high outputs, as would be expected from the hysteresis loop. Time of response for this circuit is still one half-cycle.

The response time and other characteristics of an amplifier also may be indicated by impressing a sinusoidal wave signal on the input and observing the phase shift of the output. The wave forms produced when a 10-cycle-per-second modulating frequency is applied to the input of 60-cycle carrier amplifier-modulator are shown in Fig. 20. The upper set of traces shows the signal compared with the raw output while the lower set shows the signal and the output voltage after the latter has been through a low-pass filter.

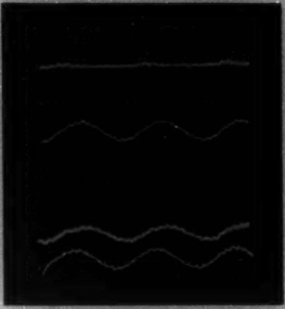


Fig. 20. Wave forms of amplifier with 10-cycle modulating signal

Top—Output voltage compared with signal voltage

Bottom—Output voltage after filtering compared with signal voltage

The oscillogram of Fig. 20 does not show clearly that the output wave form of the circuit is basically a pulse width modulation. This result occurs because the 90-degree firing-angle chopped wave form. At small signal voltages, the output pulse height reaches a value equal to the peak value of the impressed carrier voltage. The width of the pulse is determined by the volt-second content of the signal. As the signal becomes larger (the volt-second content increases), the maximum height of the output pulse remains the same, but the width increases (with a corresponding increase in energy content), until in a perfect modulator the pulse becomes half a cycle of a sinusoidal wave. This action is more clearly illustrated in the expanded oscillogram of Fig. 21, which shows approximately 1 cycle of a 12-cycle-per-second modulating voltage, and the corresponding output voltage.

In practical circuits there may be a shifting of the firing angle to slightly more or less than 90 degrees through the self-balancing feature of the circuit previously described. This shift would mean that the pulse height for the small signal would be slightly less than the peak value of the carrier voltage. Then as the signal increases, the absolute height of the output pulse would quickly increase to the peak value of the carrier voltage. This effect is apparent in Fig. 20.

The action of this circuit as a modula-

Fig. 21. Expanded sweep oscillogram of amplifier output compared with signal voltage

tor was further studied by impressing a complex wave form on the input and observing the fidelity (band-pass characteristics) in the output. By means of a Servoscope, a 40-cycle voltage was suppressed-carrier modulated at approximately a 10-cycle rate and applied to the amplifier. The amplifier input wave form therefore had sum and difference products of the two frequencies as well as the signal frequencies. This signal was applied to the amplifier with a 60-cycle carrier frequency and the output passed through a low-pass filter with a 60-cycle cutoff frequency. The output wave form is compared with the input wave form in Fig. 22. The output wave form (upper trace) was shifted physically to the left after the picture was taken to compensate for the half-cycle lag of the amplifier and allow direct comparison of the two wave forms.

The filtering problem with this amplifier modulator is lessened because the carrier frequency is inherently suppressed in the output. The main frequency component in the output wave form will be the signal frequency f_s. The next existing frequency components will be $(2f_{ac}-f_s)$ and $(2f_{ac}+f_s)$ where f_{ac} is the carrier frequency. These frequencies will usually be far enough removed from the desired signal output that the filtering problem will be simplified. There are, of course, still higher frequency components of reduced magnitude arising from the choppy wave forms of the amplifier output and frequency mixing.

The frequency response of this circuit working as a modulator is limited at lower frequencies by the volt-second capacity of the output coupling transformer and at high modulating frequencies by the carrier frequency. During tests of the amplifier with a variable modulating frequency, the output voltage (at the modulating frequency) remained constant for a constant input voltage from the lower frequency limit of the harmonic wave analyzer to the carrier frequency (60 cycles). The output then declined linearly with modulation frequency to zero at twice carrier frequency (120 cycles).

This aspect of the circuit as a modulator gains importance from the presentation by Geyger[9] of the principle of eliminating zero signal drift of servo systems from changes in amplifier components. This purpose is accomplished by utilizing an amplifier system with a high-frequency carrier modulated at motor design frequencies. The carrier-frequency components presented to the motor by amplifier component drift will have no effect upon the motor torque. Therefore, in absence of signal modulation, no drift can occur in the total system.

Previous publications[10] have stressed the desirability of using inherent demodulation characteristics of half-wave amplifiers for providing system compensation by feedback from the amplifier. An inherently demodulated signal is easily available from the amplifiers being discussed. This may be accomplished by placing two resistors in series with any two legs which gate concurrently, and using the difference voltage E_f across the two resistor ends as shown in Fig. 23. As a further refinement, resistors may also be placed in the other output legs and cross-connected as indicated by the dashed lines in Fig. 23. The signal wave form will then be a unidirectional double frequency pattern which will tend to reduce the lag in the feedback system.

Fig. 22. Filtered output wave form compared with complex signal input. Upper trace is output wave after it has been physically shifted to left. Lower trace is signal wave form

Efficiency and Installed Capacity

The question of efficiency is inextricably coupled with the question of installed capacity in the servo systems which operate on the basis of both cores saturating simultaneously at 90 degrees for no-output. This is a common mode of operation for the self-saturating doubler circuit and will be used as a comparison standard. Efficiency is defined here as the power consumed in the load at full output divided by the power supplied to the amplifier at full load.

Efficiency can be increased in these circuits by reducing the size of the current limiting resistor R_c. Reference is made to Figs. 15 and 17 which indicate no appreciable change in the response time of the circuit of Fig. 6 and of Fig. 8

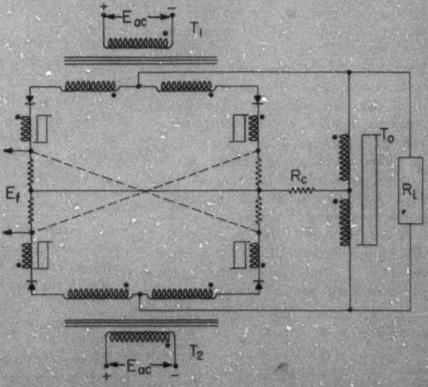


Fig. 23. Output circuit of amplifier showing take-off points for inherently demodulated feedback signal

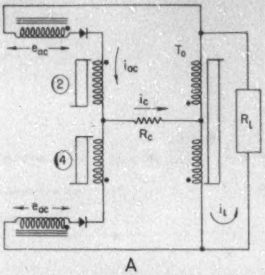


Fig. 24. Active elements in gating circuit of magnetic amplifiers

A—Gating circuit of modulator-amplifier
B—Gating circuit of saturable reactor bridge

when the compensating resistor R_c is short-circuited out. At zero output, however, the circulating currents occurring after saturation of both gating cores require high-capacity components.

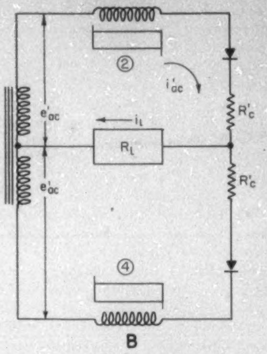
The circuit elements will have one value of average current flow at full-load operation and another sometimes greater value at no load, so the power capacities will be derived separately. The voltage and current ratings for full-load operation will be established from the specification of the load, in terms of the system efficiency. After the supply voltage has been established, it follows that the supply voltage must be absorbed over part of the gating cycle by the current-limiting resistors R_c during no-load operation. This in turn establishes no-load current ratings.

As a practical approach to this problem, the total installed capacity requirements of the circuits shown in Figs. 24(A) and 24(B) will be derived on the basis of load power requirements and desired efficiency, for no-load and full-load conditions. Total installed system capacity is defined here as the sum of the products of the voltage and current ratings of all iron-cored elements required for a complete operating system (not including prime voltage supply or load). Fig. 24(A) shows the active gating elements of the circuits presented in this paper. Fig. 24(B) shows corresponding active element of the self-saturating doubler bridge.

It can be shown, see the Appendix, that for a specified load voltage E_L, load current I_L, and full-load efficiency A, the two circuits will compare as follows:

Figure 24(A)—Installed full-load system capacity

$$W_f \leq \frac{E_L I_L(5+A)}{A} \tag{1}$$



(See paragraph in Appendix following equation 19.)

Installed no-load system capacity

$$W_n = \frac{E_L I_L\left(\frac{3}{2}+A\right)}{A(1-A)} \tag{2}$$

Figure 24(B)—Installed full-load system capacity

$$W_f' = \frac{2E_L I_L(2+A)}{A} \tag{3}$$

Installed no-load system capacity

$$W_n' = \frac{E_L I_L(3+A)}{A(1-A)} \tag{4}$$

It is obvious that the components will have to be designed to either the full-load or no-load ratings depending upon which is higher at the chosen efficiency. Superficial examination of these equations reveals that the largest difference between the two systems is in the no-load ratings which will govern in most cases. The slight difference appearing between the full-load ratings of the two systems will probably disappear after adjusting for the overrating as explained in the Appendix.

Further manipulation of these equations is dependent upon desired performance factors such as duty cycles, desired weights and efficiencies, etc. As a generalization, the ratings (component size) of the system of 24(A) will be less than that of the system of 24(B) and approach a ratio of 1 to 2 at high efficiencies. Rectifier ratings have not appeared directly in the foregoing equations but will compare as the ratings of the capacity per leg, which are derived in the Appendix. These ratings will always be smaller in the presented system.

Conclusions

The circuits shown in this presentation provide a versatile basic element in the magnetic amplifier field for servo amplifiers. This basic element possesses comparatively high gain with fast response speed and full-wave reversible output. It also has a wide range of possible sizes limited mainly by available rectifiers. This range will probably be extended greatly when some of the projected improved germanium power diodes and silicon diodes are in regular production. The basic element should be amenable to all of the various tricks promulgated in the amplifier field such as negative feedback, positive feedback, or derivative feedback. The element is susceptible of either a-c or d-c control, with fast response in either case.

The basic element seems to hold promise as a suppressed carrier modulator in circuits requiring large power outputs. A field which promises an easy immediate application is an audio amplifier with the amplifier feeding the speaker voice coil directly. The amplifier should be of use in other applications such as those requiring a reversible polarity unidirectional output with a reversible polarity a-c or d-c signal.

Appendix

Derivation of full-load capacity for the circuit of Fig. 24(A) presupposes that at full output either core 2 or core 4 would be saturated during the entire gating half-cycle and the other core must not saturate at any time during this half-cycle. Under these conditions, when A equals efficiency, the instantaneous power expression will be

$$\frac{i_L^2 R_L}{A} = \frac{i_c^2 R_c}{1-A} \tag{5}$$

if magnetizing currents are neglected, the sum of the NI on the two windings of T_0 must equal zero.

So

$$2i_L = i_c = i_{ac} \tag{6}$$

but i_{ac} current flows through the core only during the gating half cycle

average, per leg $I_{ac} = I_L$ (rated load current) (7)

From equations 5 and 6

$$\frac{i_L^2 R_L}{A} = \frac{(2i_L)^2 R_c}{1-A} \tag{8}$$

$$R_c = \frac{R_L(1-A)}{4A} \tag{9}$$

The supply voltage must equal the sum of the iR drops so

$$e_{ac} = i_c R_c + \frac{i_L R_L}{2} \tag{10}$$

From equations 6, 9, and 10

$$e_{ac} = \frac{i_L R_L(1-A)}{2A} + \frac{i_L R_L}{2} \tag{11}$$

$$e_{ac}=\frac{i_L R_L}{2A} \tag{12}$$

$$=\frac{e_L}{2A} \tag{13}$$

and average $E_{ac}=\dfrac{E_L}{2A}$ (14)

Installed full-load capacity, per leg

$$P_f=I_{ac}E_{ac} \tag{15}$$

From equations 7 and 14

$$=\frac{I_L E_L}{2A} \tag{16}$$

The output transformer T_o, must be included in the total operating system in this type of amplifier. The voltage rating of the transformer must be E_L by definition. The current rating of an autotransformer connected as shown will be I_L for full-load operation. So full-load installed capacity

$$P_{Of}=E_L I_L \tag{17}$$

The supply transformer may be wound on two cores as shown or combined on one core without affecting the total rating. During full-load operation in either case, the four supply transformer secondary windings will each be rated at E_{ac} and each winding must be able to carry an average current of I_{ac}. Neglecting transformer losses, the primary winding must be able to carry 2 E_{ac} I_{ac}.

Full-load supply transforms capacity

$$P_{Tf}=6E_{ac}I_{ac} \tag{18}$$

From equations 7, 14, and 18

$$=\frac{3E_L I_L}{A} \tag{19}$$

This expression is pessimistic as it is evident that the active windings can tolerate a current density higher than normal because of the presence of passive windings and consequent increased radiating surface area. The strict derivation of the actual capacity will not be included in this Appendix. The full-load system installed capacity will be the sum total of the ratings of all elements in the circuit.

Full-load system installed capacity

$$W_f=4P_f+P_{Of}+P_{Tf} \tag{20}$$

From equations 16, 17, 19, and 20

$$=\frac{4E_L I_L}{2A}+E_L I_L+\frac{3E_L I_L}{A} \tag{21}$$

$$=\frac{E_L I_L(A+5)}{A} \tag{1}$$

After required supply voltage has been established, in equation 14, it naturally follows from the definition of no-load operation that after 90 degrees of the gating half-cycle, the voltage must be absorbed by R_c, so

$$p_n=\frac{(e_{ac})^2}{R_c} \tag{22}$$

Current flows in each leg during one-quarter of a cycle and during this time two legs each supply one-half the current to R_c, so for each leg, installed no-load capacity, per leg

$$P_n=\frac{E_{ac}^2}{8R_c} \tag{23}$$

From equations 9, 14, and 23

$$=\frac{E_L^2}{8AR_L(1-A)} \tag{24}$$

$$=\frac{E_L I_L}{8A(1-A)} \tag{25}$$

In no-load operation, the current through each leg

$I_{ac}'=\dfrac{E_{ac}}{8R_c}$ using equation 23 (26)

From equations 9, 14, and 26

$$=\frac{E_L}{4R_L(1-A)} \tag{27}$$

Because of the circuit configuration, the output transformer must carry the sum of the average leg currents during no-load operation. So no-load output transformer capacity

$$P_{On}=4E_L I_{ac}' \tag{28}$$

From equations 27 and 28

$$=\frac{E_L I_L}{(1-A)} \tag{29}$$

The four secondary windings of the supply transformer will each have a voltage rating E_{ac} and a current rating I_{ac}'. Neglecting transformer losses the primary winding will have the same wattage rating as the sum of the active secondaries, so no-load supply transformer capacity

$$P_{Tn}=8E_{ac}I_{ac}' \tag{30}$$

From equations 14, 26, and 30

$$=\frac{E_L I_L}{A(1-A)} \tag{31}$$

The total system no-load installed capacity

$$W_n=4P_n+P_{On}+P_{Tn} \tag{32}$$

From equations 25, 29, 31, and 32

$$W_n=\frac{4E_L I_L}{8A(1-A)}+\frac{E_L I_L}{(1-A)}+\frac{E_L I_L}{A(1-A)} \tag{33}$$

$$=\frac{E_L I_L(3+2A)}{2A(1-A)} \tag{2}$$

The same method of attack will be used in deriving installed capacities for the circuit of Fig. 24(B) and, by analogous thinking, in full-load operation

$$\frac{(i_L)^2 R_L}{A}=\frac{(i_c')^2 R_c'}{1-A} \tag{34}$$

and

$$i_L=i_c'=i_{ac}' \tag{35}$$

but i_{ac}' current flows through the cores only during the gating half-cycle

so average $I_{ac}'=\dfrac{I_L}{2}$ (36)

From equations 34 and 35

$$\frac{i_L^2 R_L}{A}=\frac{i_L^2 R_c'}{(1-A)} \tag{37}$$

$$R_c'=\frac{R_L(1-A)}{A} \tag{38}$$

The supply voltage must equal the sum of the iR drops, so

$$e_{ac}'=i_L R_c'+i_L R_L \tag{39}$$

From equations 38 and 39

$$=\frac{i_L R_L(1-A)+i_L R_L}{A} \tag{40}$$

$$e_{ac}'=\frac{e_L}{A} \tag{41}$$

$$E_{ac}'=\frac{E_L}{A} \tag{42}$$

In this circuit, however, the voltage applied to the unsaturated cores will be equal to E_{ac}' plus the motor voltage. Both the gating core and the rectifier in series with the associated resetting core must be capable of withstanding this total voltage. Therefore the voltage rating of each leg must be

$$E_B=E_{ac}'+E_L \tag{43}$$

$$=\frac{E_L}{A}+E_L=\frac{E_L(1+A)}{A} \tag{44}$$

Installed full-load capacity per leg

$$P_f'=E_B I_{ac}' \tag{45}$$

From equations 36, 44, and 45

$$=\frac{E_L I_L(1+A)}{2A} \tag{46}$$

In full-load operation, the current through each half of the secondary winding will be I_{ac}' and the voltage will be E_{ac}'. This primary rating will be the same as the secondary so full-load supply transformer installed capacity

$$P_{Tf}'=4I_{ac}'E_{ac}' \tag{47}$$

From equations 36, 42, and 47

$$=\frac{2E_L I_L}{A} \tag{48}$$

Total system full-load capacity

$$W_f'=4P_f'+P_{Tf}' \tag{49}$$

From equations 46, 48, and 49

$$W_f'=\frac{4E_L I_L(1+A)}{2A}+\frac{2E_L I_L}{A} \tag{50}$$

$$=\frac{2E_L I_L(2+A)}{A} \tag{3}$$

During no-load operation, after both gating cores saturate, the limiting resistors must absorb the total supply voltage, so power in resistors

$$p_r=\frac{(2e_{ac}')^2}{2R_c'} \tag{51}$$

But current flows for only 90 degrees of each half-cycle, so average power in resistors is

$$P_r=\frac{(2E_{ac}')^2}{4R_c'} \tag{52}$$

Because each core is blocked during reset half-cycle, the average current through each core would be

$$I_{ac}'' = \frac{(2E_{ac}')^2}{8R_c'2E_{ac}'} \qquad (53)$$

From equations 38, 42, and 53

$$= \frac{I_L}{4(1-A)} \qquad (54)$$

The voltage rating of the cores is E_B, so no-load installed capacity, per leg

$$P_n' = E_B I_{ac}'' \qquad (55)$$

From equations 44, 54, and 55

$$= \frac{E_L(1+A)I_L}{4(1-A)A} \qquad (56)$$

$$P_n' = \frac{E_L I_L(1+A)}{4A(1-A)} \qquad (57)$$

During no-load operation, the current through each half of the secondary winding of the supply transformer will be the sum of the currents through two legs. The primary-winding rating will be equal to that of the secondary winding so

$$P_{Tn}' = 8E_{ac}'I_{ac}'' \qquad (58)$$

From equations 42, 53, and 58

$$= \frac{8E_L I_L}{A4R_L(1-A)} \qquad (59)$$

$$= \frac{2E_L I_L}{A(1-A)} \qquad (60)$$

The total system no-load installed capacity

$$W_n' = 4P_n' + P_{Tn}' \qquad (61)$$

From equations 57, 60, and 61

$$= \frac{4E_L I_L(1+A)}{4A(1-A)} + \frac{2E_L I_L}{A(1-A)} \qquad (62)$$

$$= \frac{E_L I_L(3+A)}{A(1-A)} \qquad (4)$$

References

1. Control of a 65-Watt 2-Phase Servomotor With a Magnetic Amplifier, **F. S. Malick.** *Electrical Engineering*, vol. 71, May 1952, pp. 466–68.

2. An Improved Magnetic Servo Amplifier, **C. W. Lufcy, A. E. Schmid, P. W. Barnhart.** *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 281–88.

3. A New Type of Magnetic Servo Amplifier, **W. A. Geyger** *Ibid.*, pp. 272–80.

4. A Fast Response Magnetic Servo Amplifier, **J. W. Kallander.** *AIEE Transactions*, vol. 74, March 1955, pp. 49–54.

5. On the Mechanics of Magnetic Amplifier Operation, **Robert A. Ramey.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

6. The Modulator Bridge, **R. K. Hellman.** *Electronics*, New York, N. Y., vol. 11, no. 3, Mar. 1938, pp. 28–30.

7. Magnetic Amplifiers of the Self-Balancing Potentiometer Type, **W. A. Geyger.** *AIEE Transactions*, vol. 71, pt. I, 1952, p. 383–94.

8. Flux-Preset, High-Speed Magnetic Amplifiers, **C. B. House.** *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 728–34.

9. Elimination of Assymetry Zero-Drift Errors in Magnetic Servo Amplifiers, **W. A. Geyger.** *AIEE Transactions*, vol. 74, pt. I, May 1955, pp. 239–44.

10. Compensation of a Magnetic Amplifier Servo System, **H. H. Woodson, A. E. Schmid, C. V. Thrower.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 8, 1952, pp. 158–65.

Selenium Rectifiers for High-Temperature Operation

G. O. ISAACSON
NONMEMBER AIEE

THE development of the selenium rectifier has provided design engineers with a reliable, maintenance-free source of d-c power. However, certain inherent thermal characteristics of the device impose many limitations on its application. Normally, the operation of the rectifier in ambient temperatures above 45 degrees centigrade (C) requires that the power output be derated to protect the unit or that additional equipment be provided for forced cooling. Operation at temperatures above 90 C severely reduces the life expectancies of the selenium cells in spite of the precautionary measures just mentioned.

Recently, however, selenium cells have been developed which are capable of full-load operation at ambient temperatures up to 125 C. With a very slight current derating, the range of application is extended to 150 C.

There is considerable theoretical information available in the literature regarding the electrical nature of selenium at these temperatures. However, from the practical engineering viewpoint, data regarding the high temperature operation of selenium cells are very limited. The purpose of this discussion is to report the results of both short-range and extended studies on the operation of high-temperature cells.

Although these cells were designed specifically for operation at elevated temperatures, in practical use they may be operated periodically under normal conditions. Therefore, before considering the performance of the cells under specific conditions, it is desirable to discuss their characteristics at various ambient temperatures. With the extensive applications now possible with these high-temperature rectifier cells, a selenium rectifier can now be operated in ambients ranging from −55 to +150 C.

It is well known that selenium, in the form used in rectifiers, has a negative coefficient of resistance, i.e., as the temperature of the material rises, the conductivity improves. Since the selenium has this property, it is important in determining the quality of a rectifier cell that the electrical characteristics be investigated for the entire temperature range in which they are to be used.

One of the most important factors to be considered in rating a rectifier is the voltage output. The power output delivered by a rectifier is directly affected by the forward resistance of the unit. It is customary in the industry to use the measure

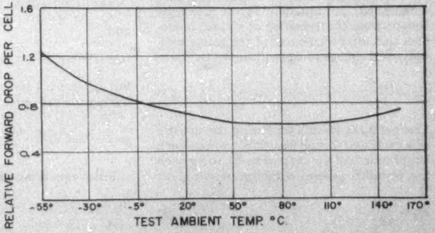


Fig. 1. Relative forward drop per cell versus ambient temperature, C

Paper 55-518, recommended by the AIEE Metallic Rectifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 25, 1955; made available for printing May 13, 1955.

G. O. Isaacson is with the Fansteel Metallurgical Corporation, North Chicago, Ill.

The author wishes to thank William E. Brown for his aid in the compilation of data used in this paper.

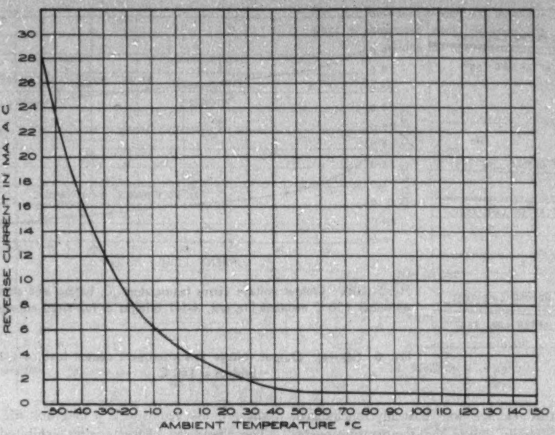


Fig. 2. Reverse current in ma alternating current versus ambient temperature, C, 4-1-1 bridges at 26 volts rms

of forward voltage drop as an indication of this resistance. (Measurements of either forward or reverse resistance with an ohmmeter have little significance.) The amount of voltage drop is obtained by passing current in the forward direction. The increase in forward resistance is therefore reflected by a rise in voltage drop.

The forward resistance characteristics of typical high-temperature cells were measured over a temperature range of −55 to 150 C with the use of the standard current density of 160 milliamperes (ma) per square inch of pulsating direct current; see Fig. 1. Average voltage drop readings of the cells indicate that a maximum of 1.2 volts occurs at −55 C. As the temperature rises, the forward drop decreases to a minimum of 0.63 volt at 90 C. From 90 to 150 C, the voltage drop increases to 0.74 volt or about 18 per cent (%) over the minimum. These voltage drop measurements indicate that the high-temperature cell exhibits the optimum forward resistance from 70 to 110 C.

Just as a temperature change affects the forward resistance of a rectifier cell, such a change would also have an effect on the reverse resistance. The reverse resistance is customarily gauged by the measure of current flow in the reverse direction as a result of voltage applied. The reverse resistance tests for the high-temperature cell, using alternating current, show an interesting pattern; see Fig. 2. As previously noted, forward voltage drop measurements indicate a maximum forward resistance at −55 C. However, at this temperature the reverse resistance is at a minimum, i.e., more reverse current is flowing at a given voltage than flows at higher temperatures. As the temperature rises, the reverse current rapidly decreases. Typical measurements for determining the reverse current flow showed 28 ma, alternating current at −55 C for 4-1-1 bridges with 26 volts rms applied; see Fig. 2. This value drops to less than 2 ma at room temperature. The reverse current remains less than 2 ma up to the maximum test temperature of 150 C. Again, it is interesting to note that optimum results were obtained at elevated temperatures.

The quantity of reverse current which results from applied alternating voltage is a widely accepted criterion of rectifier quality. However, in certain applications, the rectifier may be subjected to battery direct voltage in the reverse direction. In these instances, extensive damage to the unit may result if the reverse current is sufficient to cause overheating. Therefore, one of the most severe tests of the rectifier's blocking ability is the application of battery direct voltage. Standard test procedure in the Fansteel Laboratory calls for reverse current measurements at 22-volt battery direct current for 26-volt cells, and 28 volts for 33-volt cells. High-temperature cells were tested at various voltages while at ambient temperatures of 27, 71, and 150 C. Typical $2^{3}/_{8}$-inch diameter 26-volt high-temperature cells are normally required to show less than 20-ma reverse current at 22 volts. It should be noted that this limit is based on room temperature standards. Test samples remained below this level up to 38 volts, even on the poorest units. Measurements on the 33-volt cells were within limits at 46 volts. See Figs. 3 and 4. In both cases, the reverse current rise was less abrupt at the higher ambient temperatures than at room temperatures.

Data for this investigation serve to illustrate some of the characteristics of

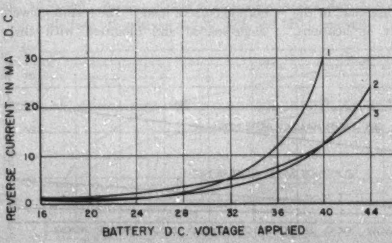


Fig. 3. Reverse current in ma direct current versus battery voltage, 26-volt type, $2^{3}/_{8}$-inch diameter cells, no. 1 at 27, no. 2 at 71, and no. 3 at 150 C

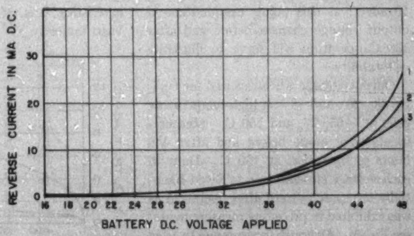


Fig. 4. Reverse current in ma direct current versus battery voltage, 33 volt $2^{3}/_{8}$-inch diameter cells, no. 1 at 27, no 2 at 71, and no. 3 at 150 C

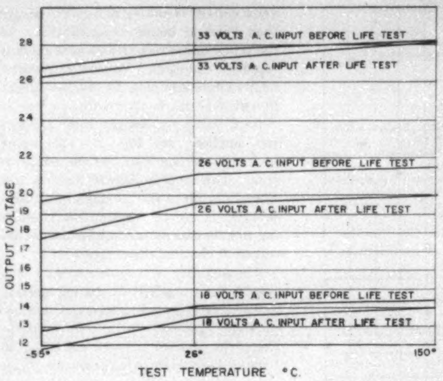


Fig. 5 (left). Output voltage versus temperature, C, before and after 456-hour 150 C ambient life test, 4-1-1 bridges at full-rated current

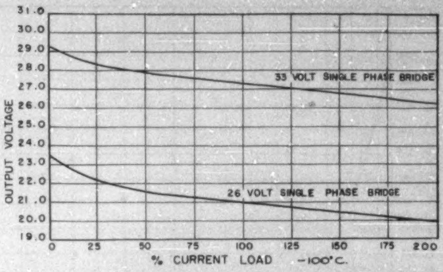


Fig. 6 (above). Output voltage versus per-cent current load, 4-1-1 bridges at 100 C

the high-temperature cell before severe operational tests. Now, the question immediately arises as to the stability of these characteristics after an extended operation at 150 C. To study the stability of the reverse resistance characteristics of the high-temperature cell, samples were tested before and after 456 hours of operation at 150 C with full-rated voltage and current. The amount of reverse current was measured at 26 volts and 18 volts rms at −55, 26, and 150 C. Under each of the test conditions the change in reverse current after severe operation was negligible.

As the discussion has indicated, the selenium rectifier has a definite normal pattern of electrical characteristics with respect to temperature changes. In addition to its immediate effects, a temperature rise breaks down or diminishes the desirable rectifier qualities. One of the most widely accepted criteria of the extent of the useful life of a rectifier is the per-cent decrease in output voltage in relation to the length of service. Later discussion will cover the progressive change in output voltage during operation. However, at this point, examination of output voltage changes before and after 150 C operation will serve to illustrate cell stability.

Output voltage was measured for typical 33-, 26-, and 18-volt high-temperature cells at −55, 27, and 150 C. Measurements were made before and after 456 hours of operation at 150 C. Data in each voltage classification followed essentially the same pattern after operation as was exhibited in preceding measurements; see Fig. 5. Although measurements after the operation indicated a decrease in outout voltage due to aging effects, the pattern of reaction of the units to temperature change was basically unchanged. Stability during high-temperature operation indicates that the unit has the basic properties of a true high-temperature rectifier. The conventional type of cell in some instances has been operated at ambients of 100 to 125 C. However, as previously mentioned, this generally required drastic reduction of power output. Severe aging and breakdown of the blocking layer still occurred after short periods of operation. In comparison, it should be noted that the high-temperature cell operates at full rated output under these conditions.

In addition to differences in the output voltage with respect to the changes in temperature, the variation at different current loads was measured. See Fig. 6. Measurements of output voltages at 100 C for typical 33-volt high-temperature cells indicated a range from 29.2 volts at no load to 26.3 volts at 200% normal load. Similar measurements on 26-volt cells gave 23.4 volts d-c output at no load and 19.9 volts at 200% normal load.

Without minimizing the necessary requirements of satisfactory electrical characteristics, it is safe to state that in the vast majority of rectifier applications, the primary requisite is a long unit life. The change in characteristics exhibited by the rectifier during use has presented one of the most perplexing design problems for rectified d-c power supplies. This change or aging of the rectifier cell must be considered in designs and compensating features incorporated in the power unit. While extensive data are available for the conventional selenium cell, little has been published to aid the designer of the rectifier power units which are used at high ambient temperatures.

In the life testing of a piece of equipment, it is customary to operate the unit under conditions simulating as closely as possible the stresses of actual operation. However, while the high-temperature cell is primarily designed to be used at high ambient temperatures, a clearer picture of the life characteristics can be obtained if the discussion begins with the results of the room temperature life tests.

Extended operation of these cells at various loads at room temperature has provided useful information about their life behavior. Typical 4-1-1 bridges were operated at one, two, and three times the normal current load. The test units were single-spaced and mounted with three

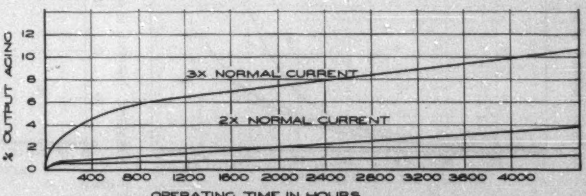


Fig. 7. Per-cent output voltage aging versus time, room temperature life tests, 4-1-1 bridges, 26 volts alternating current input

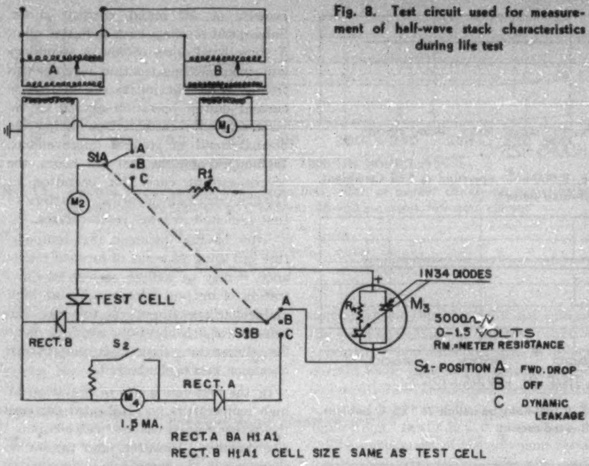


Fig. 8. Test circuit used for measurement of half-wave stack characteristics during life test

units per center bolt. Periodic measurements of output voltage were used to determine the changes in the cell conductivity. At normal load and twice normal load, the high-temperature cell shows an almost insignificant decrease in output voltage after continuous operation in excess of 5,000 hours at a 25 C ambient; average cell temperatures under these conditions were found to be approximately 30 C at normal load and 40 to 45 C at double load; see Fig. 7. At three times normal current load, the high-temperature cell can be expected to reach temperatures of 85 to 90 C. These temperatures were measured on the middle cells of the single-spaced 12-cell stack. Aging patterns of those cells operating at three times normal load illustrate one of the most pronounced peculiarities of this new cell. This phenomenon is a rapid aging in the first few hours of operation which is followed by a tendency to stabilize and remain at the new level for an extended period.

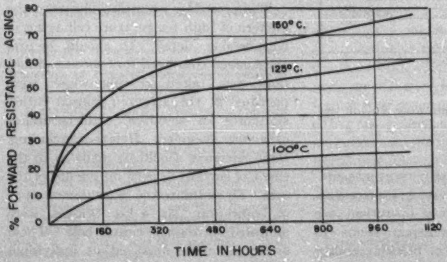
It has been found that these cells exhibit rapid initial aging if operated at normal-current load at temperatures above 100 C. This phenomenon also occurs at lower ambient temperatures if the cells are operated at higher than normal current loads. Subsequent discussion on high-temperature life tests will provide further information on the initial aging.

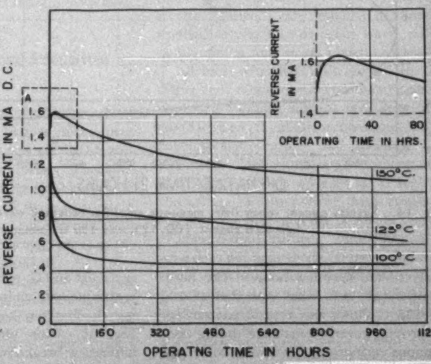
The method of determining the aging in the room temperature life tests is based on the measurement of the change in the output voltage. However, from the standpoint of classifying a metallic rectifier, a more precise test is the measurement of th eforward-resistance aging of the individual cells during half-wave operation. For test circuit, see Fig. 8. This method, in effect, determines the efficiency of the individual cell rather than giving an average performance of four or more cells, such as in life tests that measure output voltage aging. In determining the expected life of a rectifier, as based on changes in the forward-resistance half-wave single-cell stacks are operated at the ambient temperatures for which they are specifically designed. These units are tested periodically during operation to determine the reverse-current and forward-voltage drop. (Before discussing this type of life test measurement, it may be well to note that 100% increase in forward-voltage drop is roughly equivalent to 10% decrease in output voltage.) Since this method gives a more precise measurement of basic changes in the rectifier cell, the forward-voltage drop characteristics more clearly illustrate the phenomenon of high initial aging. There is a pronounced rise in forward resistance of the individual cell at 125 and 150 C ambient temperatures, while at 100 C, the cell shows a considerably less radical change in characteristics; see Fig. 9.

While the forward-resistance characteristics of the selenium rectifier undergo a unique and rapid initial change when first operated at high ambient temperatures, the reverse-resistance exhibits less dramatic behavior. During operation, the reverse current of the rectifier can be expected to diminish gradually, particu-

Fig. 9 (below). Per-cent forward resistance aging versus time, half-wave life test 26 volts rms, full-rated current at 100, 125, and 150 C ambients

Fig. 10 (right). Reverse current in milliamperes versus time, half-wave life tests, 26 volts rms at 100, 125, and 150 C ambients





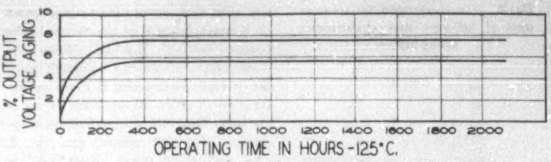


Fig. 11. Per-cent output voltage aging versus time, 2,000-hour operation at 125 C ambient, 26-volt a-c input at full-rated current

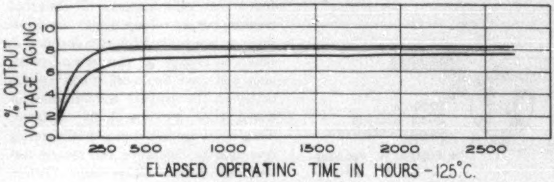


Fig. 12. Per-cent output voltage aging versus time, 2,500-hour operation at 125 C ambient, 26-volt a-c input at full-rated current

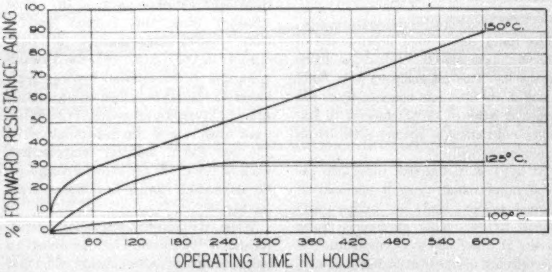


Fig. 13. Per-cent forward resistance aging versus time, magnetic-amplifier-type half-wave life tests, 26-volt rms, full-rated current, 100, 125, and 150 C ambients

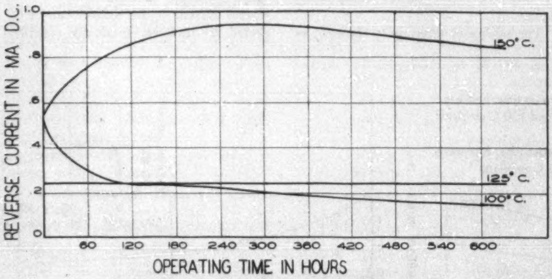


Fig. 14. Reverse current versus time, magnetic amplifier cells half-wave life tests, 26-volt rms full-rated current, 100, 125, and 150 C ambients

larly in the first few hours. The high-temperature cell, when operated at 100 or 125 C shows this typical pattern for improving reverse resistance during continuous operation. However, when operated at 150 C, reverse-current measurements indicate an increase in leakage during the first few hours of operation; see Fig. 10. This phenomenon seems to indicate a breakdown of blocking ability because of an initial thermal shock. Subsequent readings have indicated that, if some breakdown occurs, a secondary and compensating reaction reverses this trend. After the initial rise in reverse current during operation at 150 C, reverse-resistance characteristics follow the normal trend of gradual improvement. In fact, beyond the first few hours, the reverse-current curve for operation at 150 C is practically identical in pattern to that exhibited at lower temperatures.

After having discussed high-temperature cell aging in terms of forward resistance, it may be well to present an illustration of output voltage aging at high ambient temperatures. In this way, the behavior of a bridge-type selenium rectifier, during long-term high temperature operation, can be illustrated.

In the first few hours of operation at high temperature and full-rated current, the output voltage changes with comparative rapidity. However, after the initial sharp rise in the aging curve, the output voltage reaches a level of remarkable stability; see Figs. 11 and 12. In these life tests, approximately 50% of the aging occurred in the first 36 hours. Little variation was noted from 300 hours to beyond 2,000 hours of operation at 125 C.

In addition to the more routine power applications, the selenium rectifier has been called upon to perform an extremely difficult role in magnetic-amplifier circuits. Cells for these applications require exceptional stability and high reverse resistance.

The magnetic-amplifier type of cell exhibits a similar performance with much the same pattern as that of cells designed for routine high-temperature work. There is a rapid rise in forward-drop aging during the first hours of operation, but a stabilizing effect is also experienced; see Figs. 13 and 14. Reverse characteristics show the same general pattern of normal improvement at 100 C operation, namely, a stable reverse resistance at 125 C and an increase in leakage which is later reversed during operation at 150 C.

The preceding examples illustrate the pattern of high temperature cell aging in considerable detail. It should be re-emphasized that in each of the samplings, data were taken on cells which were operated at the accepted normal rating standards for selenium rectifiers, unless otherwise specified. Before concluding, some reference should be made as to the effect of current density on the life of the cell. To simulate the most severe type of operation, six 4-1-1 bridges were mounted on a single center bolt. These assemblies were operated as individual

Vol. 7
Pt. 1
1955

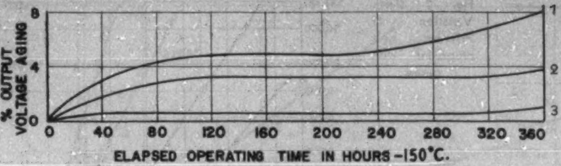


Fig. 15. Per-cent output voltage aging versus time, effect of current density on output voltage, 360-hour life tests at 150 C ambient, 26-volt a-c input, full-rated current

No. 1—Normal current density
No. 2—Two-thirds normal current density
No. 3—One-third normal current density

single-phase bridges at 150 C ambient. Two units were operated at each of the following current loads: normal, two-thirds normal and one-third normal. Each rectifier was operated at 26 volts rms; see Fig. 15.

The results of this life test indicated that aging is primarily a function of current load rather than of the ambient at which the cell was operated. It was also noted that the phenomenon of the high initial aging is to a great extent related to the current load. Those stacks operating at one-third load did not exhibit the steep aging curve in the first hours of operation. The units operating at higher current loads showed the familiar initial aging pattern. It may be concluded that a current load derating is required if extended unit life at 150 C is necessary.

The application of the selenium rectifier to high-temperature operation provides a unique test of the characteristics of selenium cells. Studies of cell characteristics indicate that selenium possesses the necessary electrical properties for high-temperature operation. Within the limits imposed by the physical properties of this semiconductor, the rectifier cells show excellent stability toward heat. Both forward and reverse resistances react favorably to thermal effects of high ambient temperatures.

Experimental evidence has indicated that neither the heat content of the surrounding cooling mediums nor the voltage applied contribute extensively to cell aging. In fact, the effect that the current load has on the aging appears to be the chief limitation imposed on this cell.

Conservatively speaking, the high-temperature rectifier offers a life expectancy of more than 2,000 hours of continuous operation at 100 C. As a matter of fact, some stacks have been in operation for more than 5,000 hours. Life expectancies at 125 and 150 C will be in excess of 300 hours when the rectifier is operated at full current load. Since the reverse current properties remain stable at high temperatures, no voltage derating is necessary. With proper derating of the current load, cells operating at 125 and 150 C will have a comparatively long life.

Heat-Resistant Polyester Magnet Wire

W. F. GILLIAM
NONMEMBER AIEE

E. M. BOLDEBUCK
NONMEMBER AIEE

J. R. ELLIOTT
ASSOCIATE MEMBER AIEE

THE INFLUENCE of chemical composition and structure on the mechanical properties and thermal stability of polymers has long been of great interest to the chemist. An outgrowth of such work in the General Electric Research Laboratory has been the synthesis of a completely organic polymer of the polyester class, a polymer shown by tests to be eminently suitable as a thin-film magnet wire insulation. Wire insulated with this polymer, which will be referred to as Alkanex* wire, possesses excellent initial properties important in manufacturing electrical equipment. Its thermal stability is such that it can be used at class *B*, 130 degrees centigrade (C), and higher operating temperatures in long-lived electric apparatus.

Preparation and Initial Properties of Alkanex Wire

Alkanex wire is generally prepared by applying the polymer to wire in the form of a solution in cresol and alkylated aromatic solvents, such as xylene. The solutions are of lower viscosity than polyvinyl formal wire enamel and may be used in either dip or die coating. The enameling operation can be performed in conventional wire-coating equipment.

The broad curing range of Alkanex enamel at a given temperature is illustrated in Fig. 1, in which a plot of several magnet wire properties as a function of wire speed is given. The data are for 50.8-mil (American Wire Gauge 16) Alkanex wire baked in an 18-foot electrically heated oven maintained at 400 C at the 12-foot level. The initial properties of wire are excellent over a speed range of 22 to 30 feet per minute.

All data presented in this paper were obtained for enamel applied to 50.8-mil copper wire, with insulation thicknesses of 1.3–1.8 mils (film increase on radius) which is equivalent to heavy-coated Formex® wire.[1] The enamel has been applied successfully to other wires, such as aluminum and nickel-plated copper.

Alkanex wire has been found to be either equivalent to or better than Formex wire in flexibility, solvent resistance, dielectric strength, abrasion resistance, cut-through temperature, and chemical resistance.[2] The cut-through temperature of Alkanex wire is at least 50 C higher than that of Formex wire; this is consistent with the proposed use of Alkanex wire at class *B* temperature. In addition, the wire does not heat shock on prewound 3X turns when heated for 10 minutes at 175 C. Extensive factory trials in many kinds of electrical equipment have further verified

---

Paper 55-494, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted February 7, 1955; made available for printing May 6, 1955.

W. F. Gilliam, E. M. Boldebuck, and J. R. Elliott are with the General Electric Research Laboratory, Schenectady, N. Y.

The authors wish to acknowledge the contributions of their associates, A. L. Marshall, M. C. Agens, F. M. Precopio, D. W. Fox, and R. E. Burnett to the synthesis of Alkanex polymer, and to thank J. W. Eustance, J. S. Axelson, F. P. Price, F. J. Lasak, A. V. White, and C. Kamienski for their contributions to this paper.

* Trade-mark applied for by the General Electric Company.

® Registered trade-mark, General Electric Company.

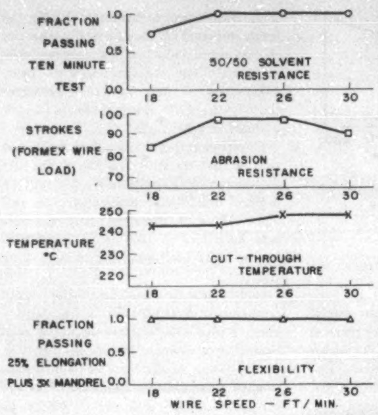


Fig. 1 (left). Variation of Alkanex wire properties with speed

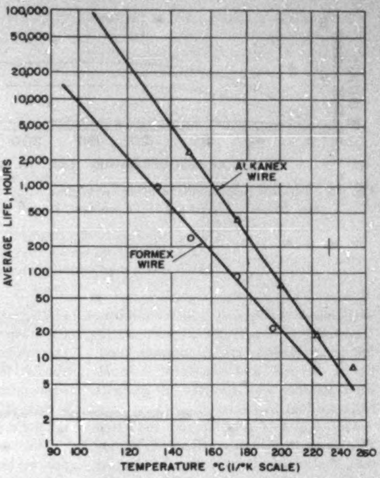


Fig. 3 (right). Weight loss. Time to 15% weight loss in circulating air

the excellent mechanical properties of Alkanex wire.

Exposure of mechanically stressed Alkanex wire for short times at room temperature to dioxane, ethylene dichloride, water, and varnish solvents such as toluene, xylene, and solvent naphtha, has no effect on dielectric strength, although slight crazing can be observed with most solvents. Under similar conditions, exposure to ethyl alcohol and diacetone alcohol lowers the dielectric strength, but heating for 10 minutes at 155 C (or 30 minutes at 135 C) returns the dielectric strength to normal values.

Thermal Stability of Alkanex Wire

Flexibility, dielectric strength, and weight-loss data as functions of time and temperature are used in this paper to describe the thermal stability of Alkanex compared to Formex wire. All aging tests were carried out on untreated wire samples suspended individually in air-circulating ovens under prevailing humidity conditions, with temperature controlled to ±3 C. Any restriction of air flow, such as by embedding the magnet wire in coil structures or by varnish treatment, markedly curtails the rate of degradation. Exposure of wire to excessive moisture, such as might be encountered in completely sealed systems containing trapped water or water-evolving components, tends to accelerate degradation at high temperatures. An accompanying paper on heat-resistant insulation systems[3] describes accelerated heat-aging tests on Alkanex wire in combination with other materials, both in functional tests and in actual equipment.

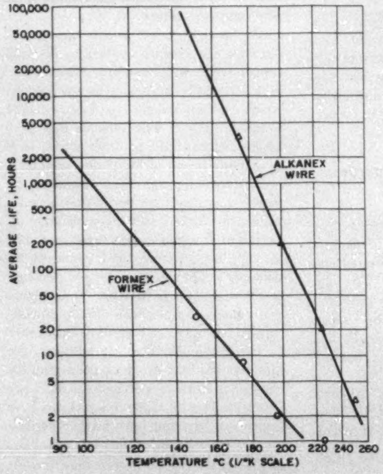


Fig. 2 (left). Flexibility retention. Time to an elongation of 30%

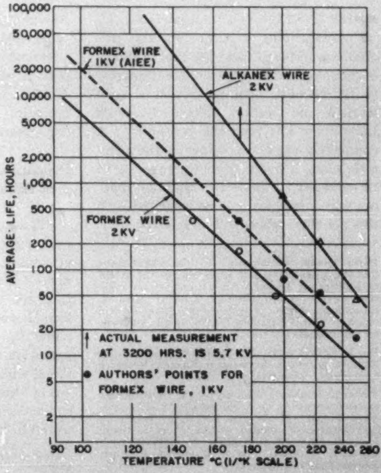


Fig. 4 (right). Dielectric strength. Time to 2-kv breakdown for unvarnished twisted pairs

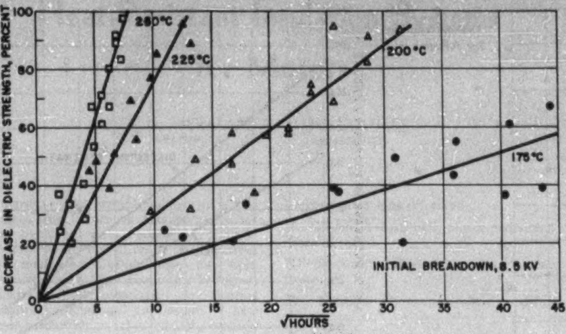


Fig. 5. Dielectric strength of Alkanex wire. Twisted pairs at various temperatures

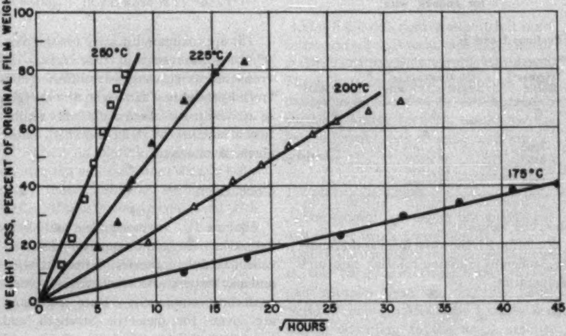


Fig. 6. Weight loss of Alkanex wire at various temperatures

l, 74
1
55

Alkanex wire maintains a high order of flexibility after long aging in air. In Fig. 2, time to decrease to an elongation of 30 per cent (%) as a function of temperature is compared with similar data for Formex wire. Extrapolation of the curve indicates that Alkanex wire will still retain an elongation of 30% after 200,000 hours at 130 C. In Fig. 3, data for time to 15% weight loss of enamel are similarly presented as a function of temperature.

The time to a 2-kv dielectric breakdown as a function of temperature is plotted in Fig. 4. These data were obtained on unvarnished twisted pairs prepared according to Army-Navy specification[4] and aged without spacers. The points indicated in Fig. 4 were estimated from smooth curve plots of average dielectric breakdown versus aging time for each temperature. In Fig. 4 the authors' values for Formex wire at a 1-kv end point are also compared with the unpublished AIEE "round robin" curve for the 1-kv thermal life of Formex wire.

An extrapolation of the data in Figs. 2 through 4 indicates that Alkanex wire has an 85 C advantage over Formex wire in retention of flexibility, a 30 C advantage in weight loss, and a 65 C advantage in dielectric strength retention.

Discussion of Dielectric Strength and Weight Loss

The dielectric strength and weight-loss data for both Alkanex and Formex wire can be reasonably approximated by straight lines passing through the origin in a plot of % decrease of the property versus square root of time as shown in Figs. 5 through 8. The slopes of these curves are plotted in Figs. 9 and 10 on a

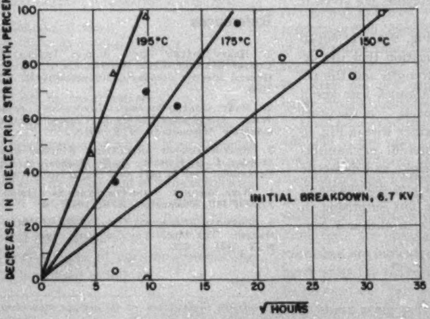


Fig. 7. Dielectric strength of Formex wire. Twisted pairs at various temperatures

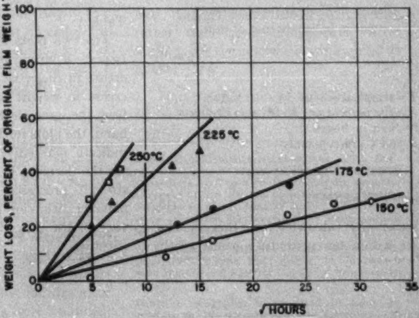


Fig. 8. Weight loss of Formex wire at various temperatures

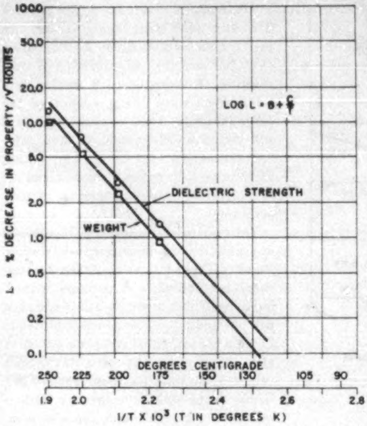


Fig. 9 (left). Thermal functions for Alkanex wire

Fig. 10 (right). Thermal functions for Formex wire

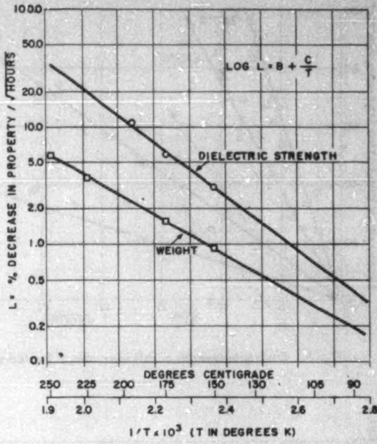


Table I. Constants of Equations 1 and 3

		Degrees Kelvin	(% Decrease in Property,/)^{2}Hours	Kcal/Mole
P, % Decrease in Property	B	C	A	E
Dielectric strength, Alkanex wire	7.3	−3200	4×10^{14}	29
Weight, Alkanex wire	7.6	−3450	16×10^{14}	32
Dielectric strength, Formex wire	5.9	−2300	60×10^{10}	21
Weight, Formex wire	4.0	−1700	0.01×10^{10}	16

log scale against $1/T$, where T is temperature in degrees Kelvin. The linearity of these plots indicates that with each polymer the mechanism does not change significantly over the greater part of the temperature range shown.

The lines in Figs. 9 and 10 may be expressed by equations of the form

$$\log_{10} L = B + C/T \qquad (1)$$

where

$$L = \frac{\text{change in \% decrease of property}}{\text{change in square root of time}} = \frac{\Delta P}{\Delta\sqrt{t}} = \frac{P}{\sqrt{t}}$$

T = temperature in degrees Kelvin
P = the percentage decrease in property
t = time in hours
B and C are constants

Equation 1 may be recast

$$\log_{10} P^2/t = 2B + 2C/T \qquad (2)$$

Values of P calculated from this equation are within ten percentage points of the observed values.

Equation 2 is the same form as the familiar Arrhenius equation relating rate and temperature

$$\log_{10} K = \log_{10} A - E/2.3RT \qquad (3)$$

where K is the rate of some process as a function of time t.[5]

The constants B and C of equation 1 may be related to the frequency factor A and the energy of activation E of equation 3; values of these constants are given in Table I.

The two plots in Fig. 9 nearly superimpose, a strong indication that the decrease in dielectric strength and the decrease in weight of Alkanex wire result from the same mechanism. On the other hand, the plots for Formex wire in Fig. 10 indicate markedly different mechanisms for the decrease in weight and dielectric strength processes. These differences in polymer behavior are not surprising in view of the differences in chemical constitution.

The approximate linearity for the plot of the square root of time versus decrease in weight suggests diffusion-controlled processes. However, the energies of activation are abnormally large for diffusion and are more in line with many chemical reactions.

Conclusion

Alkanex wire has mechanical and physical properties equal or superior to the excellent initial properties of Formex wire, and may be prepared in conventional wire-enameling equipment. Rate equations are given for dielectric strength and weight-loss data as a function of time and temperature; the thermal data indicate that Alkanex wire possesses more than adequate stability for operation at class B temperature (130 C).

References

1. Heavy Coated Vinyl Acetal Insulated Copper Magnet Wire. *Specification B22F2-S2*, General Electric Company, Schenectady, N. Y., 1953.

2. Heavy Coated Polyester Insulated Copper Wire. *Specification B22M3-S3*, General Electric Company, Schenectady, N. Y., 1955.

3. Heat-Resistant Insulation Systems for Motors, C. J. Herman, K. N. Mathes. *AIEE Transactions*, vol. 74, pt. I, 1955 (Paper No. 55-495).

4. Wire, Magnet. *Joint Army-Navy Specification JAN-W-583*, Washington, D. C., 1948.

5. Physical Chemistry (book), F. H. MacDougall. The MacMillan Company, New York, N. Y., 1937, p. 428.

Heat-Resistant Insulation Systems for Motors

C. J. HERMAN
MEMBER AIEE

K. N. MATHES
MEMBER AIEE

Synopsis: The thermal stability of a new magnet wire known as Alkanex* wire has been evaluated as a component of heat-resistant motor insulation by four different test methods. Alkanex wire has been shown to possess approximately a 40 degrees centigrade (C) superiority in thermal stability as compared to Formex ® wire when used with suitable varnish systems. It is important to choose varnish and solvent systems which are compatible with the wire enamel. It is concluded by test that Alkanex wire can be used in equipment operating at class *B* temperature and that use need not await confirmation by long service trials.

l, 74
1
55

A NEW MAGNET wire enamel known as Alkanex has been developed[1] which, because of its excellent windability and heat resistance, will have wide application in the insulation systems of a large variety of electric equipment. In this paper the use of Alkanex wire in heat-resistant motor insulation systems is discussed. Alkanex wire is compared with several other magnet wires. The results of four different types of tests have been used to determine thermal stability in combination with several different types of varnishes.

Temperature Stability of Insulation Systems

Temperature stability has been investigated by several methods which will be divided in this presentation as follows:

1. Functional tests
   a. Varnished twisted pairs.[2]
   b. Deflected beam film rupture tests.[3]
   c. Motorette tests.[4]
2. Tests on small motors

In the data, Alkanex wire, labeled *A*, has been compared with a wire of similar build and mechanical properties, heavy Formex wire (HF), and in some cases with heavy Formex glass wire and several other new magnet wire insulations labeled here as *B*, *C*, and *D*.

Several types of varnish systems also have been used as follows:

I—Pigmented long-oil alkyd.

II—Oil-modified phenolic plus a pigmented short-oil alkyd.

III—Silicone, a moderately soft, flexible type.

IV. Modified silicone, an alkyd modified silicone.

V—Modified polyester solution.

VI. Epoxy type.

VII—Oil-modified phenolic.

VIII—Modified urea phenolic alkyd.

IX—Unmodified phenolic.

This list is descriptive although it must be recognized that each type may include materials with widely varying properties. The important general problem of compatibility is discussed in the Appendix. In the following sections, results for the four different types of aging tests are described.

Twisted Pair Tests

The twisted pair aging test makes use of a simple twisted wire sample. In contrast to the untreated twisted pair samples in the companion paper,[1] the twisted pair samples described in this paper were all impregnated with varnish and baked, according to the usual practice used with actual machines. The varnished samples are oven-aged and removed at intervals for voltage test. In making twisted pair tests, the sample preparation technique as developed in the AIEE was followed and the AIEE test procedure used, except that an end point of 2 kv rather than the specified 1 kv was used. (Experience has indicated that 1 kv often did not cause failure even in badly cracked samples.)

Tests have been made at 160, 180, 200, and 240 C and many of the tests at the lower temperatures are still in progress. Results in Table I provide a basis for comparison between wire enamels and varnish systems, which will be amplified as more data are obtained. Differences in varnishes are indicated. Based on average results the polyester varnish appears to be excellent, but erratic failures after very short aging periods were obtained, particularly with Formex wire and to some extent with the other wires. In Fig. 1 Alkanex wire is compared to the curve developed for Formvar enameled wire and a temperature advantage of about 40 C over Formvar enamel is indicated.

Deflected Beam Film Rupture

The deflected beam film rupture test[3] utilizes a varnished bundle of wires which is deflected at the center until film rupture (cracking) occurs. A 500-volt check test is used to determine when cracking has occurred.

The deflected beam film rupture test is particularly useful because, unlike the twisted pair and the motorette test, the progress of thermal degradation can be followed by the changes in elongation to film rupture during the aging test. The calculated values of elongations to film rupture for the different wires with three varnish systems are plotted as a function of aging time, at 200 C in Fig. 2 and at 240 C in Fig. 3. At 200 C, 1,900 hours, (about 80 days) is not long enough to produce failure with either wire *A* (Alkanex) or wire *D*, but total failure of HF and wire *B* occurs in from 300 to 500 hours. In Fig. 2, failure of HF and wire *B* occurs so rapidly that results are not plotted. At 240 C, the behavior of wire *A* (Alkanex) and wire *D* are particularly interesting. Wires with varnish system II (oil-modified phenolic and pigmented short-oil alkyd) first lose flexibility but then increase in flexibility before a final decrease.

Motorettes

Motorette tests are advantageous where size or other problems limit the testing of actual equipment.[4] The motorettes were made and evaluated in accordance with AIEE Standard *1C*.[5] In every case a silicone-pasted mica slot liner, silicone-treated sleeving for lead and crossover insulation, and silicone glass top sticks were used to insure failure of the wire rather than other components. After initial tests, the motorettes were cycled as follows:

1. Heat aging 2 days at 240 C, 2 days at 200 C, and 8 days at 160 C.

---

* Trade-mark applied for by the General Electric Company.

® Registered trade-mark, General Electric Company.

Paper 55-495, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted February 7, 1955; made available for printing May 11, 1955.

C. J. Herman is with the General Electric Company, Fort Wayne, Ind., and K. N. Mathes is with the General Electric Company, Schenectady, N. Y.

Grateful recognition is expressed to many co-workers in the General Electric Company who made this paper possible. Particular appreciation is expressed to F. J. Lasak, H. I. Morgan, and R. L. Balke.

Table I. Life in Hours, Twisted Pair Samples, 2-Kv Tests

Varnish Treatment	Aging Temperature, C	HF			Wire A, Alkanex			Wire B			Wire D		
		PA*	5%†	95%†	PA	5%	95%	PA	5%	95%	PA	5%	95%
Oil-modified phenolic and pigmented short-oil alkyd, II	180	850	630	1,070					1,170‡				
	200	320	130	515	1,640	1,060	2,220	1,080	810	1,350‡	1,830	1,250	2,400
	240	74	54	93	350	220	480				580	415	745
Silicone, III	160	2,100	1,520	2,660					1,030‡				
	180	425	310	540				555	455	650			
	200	< 80				>1,516		150	25	275		>1,516	
	240				280	145	420	<24			595	540	650
Modified silicone, IV	180	<570						765	615	910			
	200	170	125	215		>1,516		<80				1,420	
	240	<28			170	115	230					<49	
Modified polyester, V	180	895	§					1,000	725	1,270			
	200	775	§			>1,516			>992			>1,516	
	240	72.5	§		180	170	190				240	85	305
Epoxy type, VI	200	510	460	560	1,770	580	2,960				2,100	1,560	2,620
	240	<120			205	170	240				285	215	355

* Probable average.

† 5% and 95% of values are determined from arithmetic probability paper.

‡ Point not adequately established by data available.

< = less than. Data not definitive for lower values.

> = greater than. Tests still in progress.

§ Results are skewed; not a normal distribution.

2. Vibration.
3. Moisture, 3 days at 100 per cent (%) relative humidity, 33 C.

The 120-volt test between turns was applied after each part of the cycle. With very few exceptions failure occurred after a moisture cycle.

A typical pattern of failures for several wires treated with varnish system II (oil-modified phenolic and pigmented short-oil alkyd) is given in Fig. 4. A temperature-life curve for motorettes is plotted in Fig. 5; the curve for HF is taken from extensive data presented by Gair[6] which is well established and in good agreement with data obtained in this program. Wire *A* appears to possess about a 35 C temperature advantage over HF based on motorette data.

MOTOR TESTS

Since 1952, over 200 accelerated life tests have been made on small motors to measure performance of the various insulation systems substantially in accordance with AIEE Standard No. *1C*. One of the better ground systems conventionally used in class-*A* motors was used also in the tests on heat-resistant wire insulation because it had been shown to have superior heat resistance and excellent mechanical properties. Experience has shown that there are great differences in the compatibility of various varnish treatments with the several magnet wires and ground systems.

To determine the merit of a new magnet wire, it may be helpful first to study the magnet wire characteristics desired by the motor designer as follows:

Aids to the manufacture of motors are

1. Windability by machine
   a. Abrasion resistance.
   b. Withstand elongation
   c. Limpness.
   d. Slipperiness.
2. High continuity.
3. Varnish compatibility.
4. Solvent resistance.
5. Good shelf life.

Aids to the design and usage of motors are

1. Heat resistance.
2. Sustained voltage endurance.
3. Small build, permitting smaller devices.
4. Low cost.
5. Resistance to diverse environments.

Obviously, many of the properties which are listed as aids to the manufacturing of motors are also aids to the design and usage of motors, and vice versa.

Windability by machines has been established by carefully controlled production winding trials on several types of winding machines. Physical toughness coupled with limpness and slipperiness are prime requirements for magnet wire used in small motors. For economical manufacturing most of these devices are

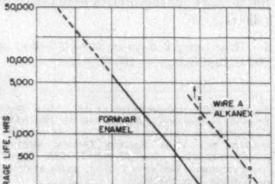


Fig. 1 (left). Twisted pair test, aging versus temperature

Varnish II, oil-modified phenolic plus a pigmented short-oil alkyd
Varnish III, silicone, a moderately soft, flexible type

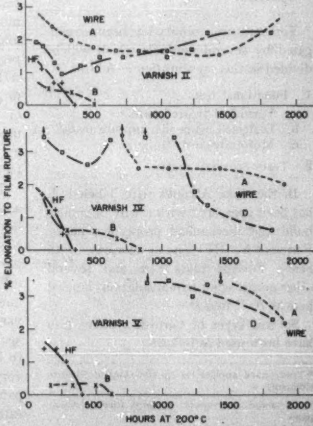


Fig. 2 (right). Deflected beam film rupture test, aging at 200 C

Varnish II, oil-modified phenolic plus a pigmented short-oil alkyd
Varnish IV, modified silicone, an alkyd-modified silicone
Varnish V, modified polyester solution

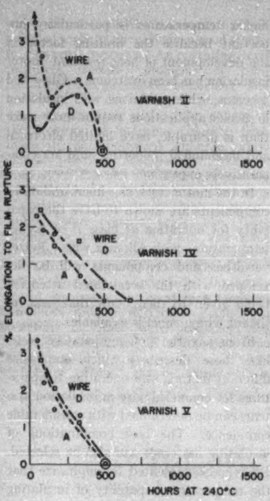


Fig. 3 (left). Deflected beam film rupture test, aging at 240 C

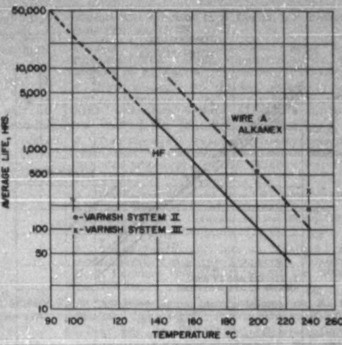


Fig. 5 (right). Motorette test, aging versus temperature

Varnish II, oil-modified phenolic plus a pigmented short-oil alkyd
Varnish III, silicone, a moderately soft, flexible type

wound on automatic or semiautomatic winding machines at high speeds. Often the wire enamel must withstand high abrasion and elongation stresses. Laboratory measurements on the wire indicate only approximately its suitability for machine winding processes. Comparative windability by machines has been studied on magnet wires *A*, *B*, *C*, and *D* with the use of Formex wire as the production standard for comparison.

Alkanex wire has proved to be at least comparable to Formex wire in all the properties listed as aids to the manufacture of motors. Other new wire enamels *B*, *C*, and *D* have performed well, but have shown some weaknesses as regards one or more of the following properties: abrasion resistance, ability to withstand elongation, or shelf life.

Heat resistance, sustained voltage endurance, and small build are the outstanding properties of several new magnet wires. The heat resistance and sustained voltage endurance have been measured on several types of small motors with various varnish treatments in accelerated life tests performed substantially as specified in AIEE Standard No. *1C*. Data indicate that the choice of varnish treatment is extremely important.

Fig. 6 shows corona starting voltage between start and run windings versus continuous accelerated life in hours for typical resistance split-phase small motors. Five motors with Formex wire are compared with five motors with Alkanex wire during accelerated life test according to AIEE Standard No. *1C*. The motors were identical except as follows: Formex wire units had heavy Formex glass wire on the start windings and heavy Formex wire on the running windings, whereas Alkanex wire units had heavy Alkanex wire on both start and run wire windings. Fig. 6 shows the remarkable sustained voltage endurance property of Alkanex wire compared to Formex wire, as measured by the sustained high value of the corona starting voltage. The small dispersion in the corona start voltage of the Alkanex wire-wound motors as compared with the Formex wire-wound motors illustrates another good characteristic of the Alkanex wire.

Relatively small dispersion of times to failure appears in the accelerated life test results for Alkanex wire-wound motors. Corona start voltage continues to be a useful measurement to indicate degradation of the insulation system as reported previously.[7] It is interesting that the two Alkanex wire-wound units which failed between windings showed lowered corona start voltage readings within 10 to 45 hours before actual failure of the motors.

The life expectance curve of insulation systems using Alkanex wire shown in Fig. 7 indicates that, for equivalent life, Alkanex wire is approximately 40 C better than the mean values for the curves I,[7] II,[8] III,[9] previously reported for Formvar insulated wire. In general, magnet wires *B*, *C*, and *D* have been life-tested at only the single temperatures, of either 160 or 180 C. In summary, these limited tests indicate that Alkanex wire and type *D* are fully comparable thermally

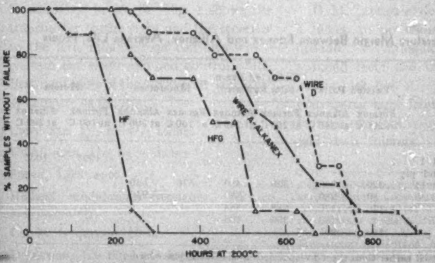


Fig. 4. Motorette test, aging at 200 C

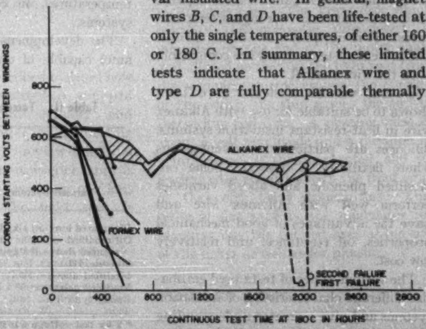


Fig. 6. Motor tests corona start voltage versus aging at 180 C

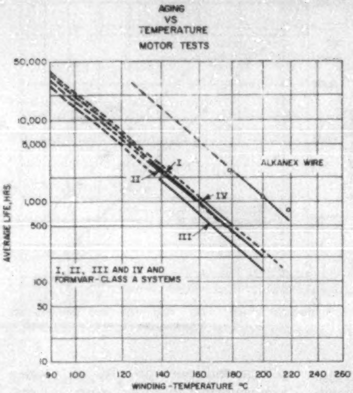


Fig. 7. Motor tests aging versus temperature

I—Reference 7
II—Reference 8
III—Reference 9
IV—Previously unpublished data per AIEE Standard No. 1C

and as regards voltage: that type *C* is a good second-place enamel in these respects; while type *B* appears to be in third place by a substantial difference.

Summary

A summary comparison of Formex wire and Alkanex wire is given in Table II. Formex wire is compared with Alkanex wire using results from four different types of test. To add significance to the comparison, results are reported for Alkanex wire aged at 40 C higher temperature than the temperature at which the Formex wire has been aged.

From the results, it is apparent that the temperature stability of Alkanex wire is at least 40 C greater than for Formex wire. Some of the varnish systems studied show a degree of superiority with a particular wire but these relatively minor variations do not affect the validity of the over-all observation that Alkanex wire is capable of operation at class *B* or higher temperatures. It is important to choose varnish and solvent systems which are compatible with the wire enamel.

Several varnish systems have been shown to be suitable for use with Alkanex wire in heat-resistant insulation systems. Silicones are particularly advantageous where flexibility is desired. Some oil-modified phenolic and alkyd varnishes perform well with Alkanex wire and have the advantages of good mechanical properties, oil resistance, and relatively low cost.

The various types of tests used emphasize different characteristics of insulation systems and give a greater understanding of the changes that occur during aging than would a single type of test. The deflected beam film rupture test emphasizes flexibility. The motorette tests involve vibration and severe moisture cycles. Tests on small motors are particularly significant since they involve the operation of actual equipment, but small motor tests alone may not always correlate with the performance of larger equipment. The results of the four different types of aging test strongly support the suitability of Alkanex wire for operation at class *B* or higher temperatures.

Conclusions

Alkanex wire has been shown by four different types of temperature-aging test to possess approximately a 40 C superiority in thermal stability as compared to Formex wire when used in combination with several types of varnish systems. Thus, Alkanex wire is suitable for operation at class *B* (130 C) or somewhat higher temperatures, in compatible insulation systems.

The development of organic wire enamel capable of operating at class *B* or higher temperatures is particularly important because the limiting factor in the development of heat-resistant motor insulation has been overcome. Glass and asbestos, while useful as wire insulation in specific applications, require more space than is desirable, have limited electrical and mechanical properties, and are comparatively expensive.

In the motor tests, slot liners and other components are shown to have the capability for operation at class *B* or higher temperatures with Alkanex wire. Better varnishes and components will be developed with the accelerated attention being given to them, now that heat-resistant wire enamel is available.

From a variety of comparative tests, like those described which emphasize different factors, safe operating temperatures for enameled wire in insulation systems can be determined with considerable confidence. The best combinations of insulating materials can also be selected. Evidence is presented to emphasize that the temperature capability of insulating materials may be determined by test rather than by definition alone. The confirmation by actual experience as it is attained is, of course, important but general acceptance of new materials like Alkanex wire need no longer await long, limited, and hesitant service trials of many years' duration.

Appendix. Compatibility of Materials in High-Temperature Insulation Systems

The importance of evaluating insulating materials in combination as used has long been recognized. Combinations of magnet wire with several varnishes are described in the body of the paper. Some more general observations are made here.

Many available materials may have particular suitability for operation at class-*B* temperatures and the particular choice depends upon the specific conditions of use. Silicones, oil-modified phenolics and alkyds,

Table II. Temperature Margin Between Formex and Alkanex, Average Life, Hours

	Twisted Pairs*		Deflected Beam Film Rupture†		Motorettes		Motors	
Varnish System	Formex at 200 C	Alkanex at 240 C	Formex at 200 C	Alkanex at 240 C	Formex at 160 C	Alkanex at 200 C	Formex at 160 C	Alkanex at 200 C
Pigmented long-oil alkyd, I							1095	1035
Oil-modified phenolic and pigmented short-oil alkyd, II	320	350	300	480	570	540		
Silicone, III	80	280		886		940		
Modified silicone, IV	170	170	288	384				
Modified polyester, V	‡	180	360	408				
Epoxy type, VI	510	205						

* 2-kv test voltage was used rather than the 1-kv specified in the AIEE procedure.

† End point chosen at 1/2% elongation.

‡ Average is invalid because of many early failures resulting in pronounced skew.

alkyd-silicones, alkyd-melamines, and oil-modified asphalts have all given good results in particular systems. As operating temperatures increase, problems of compatibility become more important and involved. For example, at temperatures over approximately 105 C the danger of the evolution of hydrogen chloride from materials like polyvinyl chloride, vinyl chloride plastisols, neoprene, etc., increases markedly. Contact with hydrogen chloride will degrade and rapidly embrittle many materials, particularly at high temperatures. Some organic acids like those present or produced in rosin and some modified asphalt potting compounds may also result in degradation at high temperatures. Moisture, particularly in tightly sealed systems operating at high temperatures can produce degradation, and even cellulosic or other water-producing materials may need to be eliminated from tightly sealed systems.

Solvent effects must be considered in the application of varnishes, and the result of solvent attack may become particularly pronounced in operation at high temperatures if solvents are trapped in the curing process. Styrene, a component of many but not all polyester resins, may, for example, have a pronounced solvent action on many wire enamels. The only sure approach is to make adequate tests on proposed combinations. Such tests should be made at several temperatures and for an adequate period of time.

References

1. HEAT-RESISTANT POLYESTER MAGNET WIRE, **W. F. Gilliam, E. M. Boldebuck, J. R. Elliott.** *AIEE Transactions*, vol. 74, pt. I, Sept. 1955, pp. 557–60.

2. THERMAL EVALUATION OF INSULATING MATERIALS. *AIEE Special Publication S-61*, "A Method for Evaluation of Thermal Stability for Magnet-Wire Enamel," **C. G. Currin, J. F. Dexter.** June 1954, pp. 19–25.

3. *Ibid.*, "Film Rupture Versus Aging for Enameled Wire Varnish Combinations, **H. I. Morgan, K. N. Mathes.** Pp. 26–30.

4. NEW PERFORMANCE STANDARDS FOR ELECTRICAL INSULATION OF ROTATING MACHINES, **AIEE Committee Report.** *AIEE Transactions*, vol. 73, pt. III-B, Dec. 1954, pp. 1542–46.

5. TEST CODE FOR EVALUATION OF SYSTEMS OF INSULATING MATERIALS FOR RANDOM-WOUND ELECTRIC-MACHINERY. *AIEE Standard No. 1C*, Jan. 1954.

6. INSULATION SYSTEMS FOR RANDOM-WOUND MOTORS EVALUATED BY MOTORETTE TESTS, **T. J. Gair.** *AIEE Transactions*, vol 73, pt. III, 1954 (Feb. 1955 section), pp. 1702–07.

7. MOTOR INSULATION LIFE AS MEASURED BY ACCELERATED TESTS AND DIELECTRIC FATIGUE, **C. J. Herman.** *Ibid.*, vol. 72, pt. III, Oct. 1953, pp. 986–91.

8. AGING OF SMALL MOTOR INSULATION, **K. N. Mathes.** *Ibid.*, vol. 71, pt. III, 1952, pp. 254–62.

9. A METHOD OF EVALUATING INSULATION SYSTEMS IN MOTORS, **C. B. Leape, J. McDonald, G. P. Gibson.** *Ibid.*, vol. 72, pt. III, 1953, pp. 793–99.

Ticketfax Electronics

C. JELINEK, JR.
ASSOCIATE MEMBER AIEE

TICKETFAX is a special-purpose facsimile equipment developed by the Western Union Telegraph Company and used primarily for the reproduction of the intelligence from a specific size of copy. It is used by the Pennsylvania Railroad in Philadelphia and the New York Central at Cleveland to facilitate the handling of Pullman space accommodations.

Modern systems of reservation-handling require high-speed communications. Wherever transmission facilities permit, 1,800-rpm Ticketfax equipment is used; this, to the communications engineer, represents an area speed equivalent to 700 words per minute if elite type is used. On longer circuits, or on those with restricted bandwidth, 360-rpm machines are used. Because of some novel features of the 1,800-rpm equipment, it will be described in detail. Fig. 1 shows the transmitter to the right and the recorder to the left, mounted on consoles which contain the electronic and control units. Pertinent facts of the 1,800-rpm Ticketfax system are as follows:

Copy size, $1^{1}/_{2}$ by $4^{1}/_{2}$ inches.

Useful area, $1^{1}/_{2}$ by 4 inches.

Speed, 1,800 rpm.

Highest modulation frequency, 7.2 kc approximately.

Carrier frequency, 10 kc.

Definition, 120 lines per inch.

The modulation frequency range is from a fraction of 1 cycle per second to 7.2 kc, i.e., an 0.008-inch spot or stylus can resolve approximately 60 alternate black and white lines, or 60 cycles per inch by 4 inches by 30 revolutions per second equals 7.2 kc. The lower limit is derived from scanning a black area for, say, 15 revolutions and a white area for 15 revolutions to generate a frequency of 1 cycle per second. At 120 lines per inch definition and 30 revolutions per second, the lineal scanning rate is 1/4 inch per second, so that at $1^{1}/_{2}$-inch wide ticket is scanned in 6 seconds.[1] With a carrier frequency of 10 kc and modulation on a double-side-band basis, a bandwidth of at least 14 kc is required, extending from 3 to 17 kc.

Transmitter Functioning

D. M. Zabriskie[1] described in detail how a ticket to be transmitted is wrapped around a transparent cylinder and scanned from the inside by a rotating mirror, see Fig. 2. To review briefly, a converging light from an exciter lamp is focused onto the copy by means of a lens and two mirrors. The light spot is rotated around the inside of this cylinder of copy by one of the mirrors which is mounted on the end of the motor shaft. Reflected light is picked up by this same rotating mirror, brought out of the end of the cylinder, and focused onto the cathode of a photomultiplier by means of a Fresnel lens.

The photomultiplier is a conventional *931* and light signals (variations of reflected light from the copy being scanned) are amplified to cause a voltage swing as high as 20 to 30 volts across the anode load resistor. A cathode follower is used to bring these signals down to the modulator on a low-impedance basis. A preamplifier chassis containing a small *RF*-type high-voltage power supply for the photomultiplier, and the cathode follower is mounted in the transmitter proper.

Phasing pulses which can be called horizontal framing pulses are generated by a magnetic slug traversing a gap in a pickup coil core. The brass disk which contains the magnetic slug insert is mounted on one end of the motor shaft and aligned so that, as the light spot from the scanning mirror crosses the overlap of the paper, the slug passes the gap in the pickup coil core. A phasing pulse of about 5-volt amplitude and a duration of 1 millisecond is thus generated, and its leading edge coincides with the front edge of the copy being transmitted. This pulse is also used in the modulator for blanking out the colored area at the end of the ticket as well as for phasing.

Modulator Action

The modulator chassis contains the carrier oscillator, balanced modulator

---

Paper 55-502, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 29, 1955; made available for printing May 11, 1955.

C. JELINEK, JR., is with the Western Union Telegraph Company, New York, N. Y.

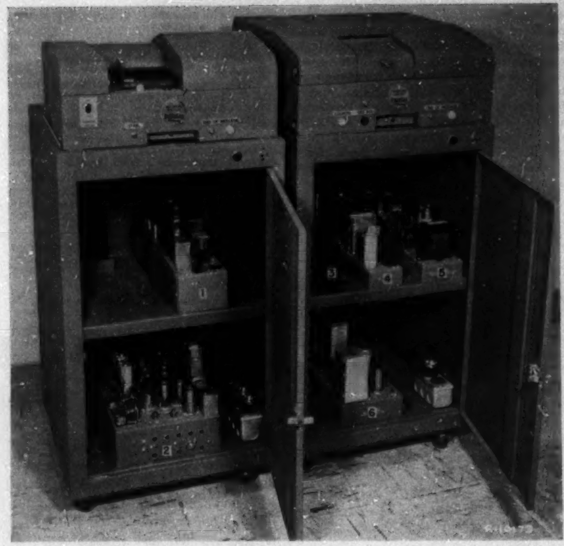


Fig. 1. Ticketfax transmitter and recorder consoles showing placement of electronic and control units

1—Recording amplifier
2—Recorder control unit
3—Transmitter exciter lamp transformer
4—Low-voltage regulated power supply
5—Transmitter control unit
6—Modulator chassis

and amplifier circuits, the blanking pulse amplifier, the multivibrator, and the changeover relay. The oscillator and the accompanying filter and line equalizer are plug-in units, making the modulator a basic unit which can be used with facsimile apparatus operating at other speeds.

The circuits operate as follows: As soon as the message is wrapped and the scanning motor is running, signals and phasing pulses appear on the contacts of the changeover relay. Now, the signal contacts are normally open so that before the changeover relay operates only the phasing pulses appear at the input to the modulator diodes. The pulses, as derived from the pickup coil, are first amplified and inverted and then used to trigger a monostable or one-shot multivibrator. This multivibrator stage serves to sharpen the edges of the pulse and also, because it contains an adjustable resistance-capacitance circuit, to make the same modulator chassis usable when different blanking areas are required and/or when a longer phasing pulse is needed, as in the 360-rpm equipment. The positive pulses on the anodes of the modulator diodes cause the diodes to conduct and thus provide a burst of carrier to the amplifier stages and thence to the line each time the magnetic slug passes the gap in the pickup coil core.

When the changeover relay operates, phasing pulses are removed from the input to the modulator diodes and signals generated by the intelligence scanned are substituted. At the same time another pair of contacts connects the phasing pulses to the cathode side of the modulator. Now, when positive phasing pulses are connected to the cathodes of the modulator diodes and if these pulses are equal to or greater in amplitude than the pulses of intelligence which are on the anodes of the diodes, no modulated carrier can appear at the output, and any intelligence scanned during the phasing-pulse duration is said to be blanked.

To insure freedom from drift and to limit unbalance and frequency discrimination caused by distributed capacitance, etc., a low-impedance modulator is used. Clamping to background or d-c restoration is accomplished in the grid side of the cathode follower feeding the modulator, thereby providing a long-time discharge path for the coupling capacitor. The charging time, however, is very short, since the charging current is taken from the output of the cathode follower in the preamplifier. This circuitry insures proper clamping even at low scanning speeds.

To provide a high degree of linearity for both large and small signal input levels, the modulator diodes are operated at high bias and signal amplitudes. This type of operation causes considerable third and higher order harmonics to be generated; therefore a filter is used in the output. The filter is a "band-pass" type, thus it also attenuates modulation frequencies which may be present in the output because of modulator unbalance. With the use of matched diodes and sensitive instruments for balancing procedures, 50-decibel ratios of modulation products to modulation frequencies in the output have been attained. Balance ratios in the average production model are of the order of 40 to 45 decibels.

The output of the modulator filter is fed into an attenuation equalizer and from there to the line through two stages of am-

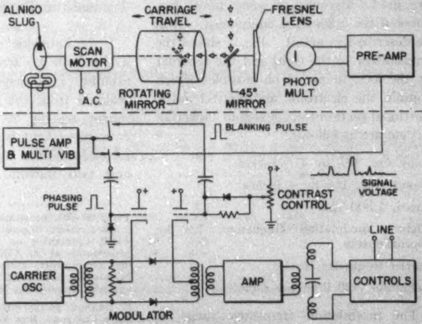


Fig. 2. Transmitter simplified schematic

plification. Because most of the lines used by the railroads are large conductor pairs, amplitude equalization is not difficult even though the transmitted frequency band is from 3 to 17 kc.

Phasing Theory

Before describing in detail how the receiver operates and how the phasing pulses are used to phase, or horizontally frame the message, the phenomenon of phase position of the armature of the synchronous motor should be discussed.

The armature of a 4-pole 1,800-rpm salient-pole synchronous motor has four possible phase positions; i.e., if the angular position of a mark on the shaft of this motor is observed with a stroboscope locked to the 60-cycle line frequency, the mark will appear in one of four positions each time the motor is started. These positions are 90 degrees apart, corresponding to the four poles, and if the line voltage and load conditions remain the same each time the motor is turned on, this phase position is maintained with a high degree of accuracy. By using the Alnico instead of the mark on the motor shaft, and by angularly positioning this slug in respect to a particular angle of current flow through the motor, the scanning motors used in this system are aligned so that the phase positions of their armatures are identical.

One way of making this alignment, Fig. 3, is to connect the vertical deflection plates of an oscilloscope to a voltage taken from across the 60-cycle line feeding the motor connected in series with the output of the pulse coil. The presentation will show the pulse at a particular angle of current flow through the motor. The Alnico slugs on the shafts of all scanning motors are aligned in this way and pinned to the shafts.

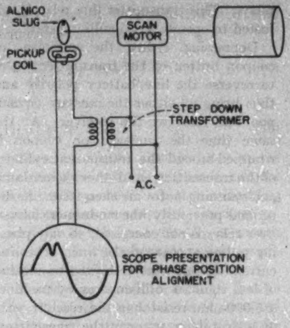


Fig. 3. Method of phase-position alignment

Recorder Synchronization

Of course only one of the phase positions of any motor armature is correct in respect to the phase position of a second and duplicate motor. Now if one motor, e.g., the recorder of Fig. 4, is momentarily interrupted so as to cause it to slip one pole at a time while the transmitter motor continues to run synchronously, a condition will be established where the magnet slugs on both shafts will be at the exact same angular position. At this time, the pulses derived from the pulse pickup coils will be in coincidence and if no further interruption of the motor circuit is made, the recorded message will be properly framed.

This is the method used to phase the recorder to the transmitter in the Ticketfax system. Actually, an adjustable-width motor-driven commutator is used to interrupt the recorder stylus motor. For proper pole slippage it was found that the stylus motor can be interrupted for a period of approximately 35 milliseconds. Since the motor-driven commutator rotates twice per second, and the scanning motor 30 times per second, as many as 15 coincidence pulses may be generated during the condition when the magnets on both shafts are in the same angular position.

At the receiver the phasing pulses coming in from a remote transmitter are first amplified and detected; a filter removes most of the carrier frequency, thus a negative going pulse of about 5- to 6-volt amplitude is provided at one control grid of a coincidence tube. The second control grid of this coincidence tube is connected to the local pulse coil. The tube circuit values are such that neither pulse alone can bias the tube to cutoff but, when the grid pulses coincide, no plate current can flow and the plate voltage rises to very nearly the $B+$ value. A diode is connected to the plate of this coincidence tube and so biased that no conduction can take place until the high coincidence voltage occurs, at which time a large voltage drop is present across the diode load resistor. The coincidence pulse thus generated has an amplitude of about 70 volts, a duration determined by the narrower of the two pulses from which it was derived, and a repetition rate of 30 per second.

To make sure that the coincidence pulse is not caused by a severe line hit or a fortuitous ouput at a time when the recorder motor is about to be interrupted, the coincidence pulses are integrated and a portion of this delayed voltage is used to trigger the phasing relay thyratron. A

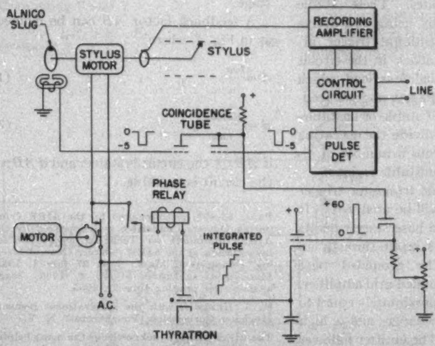


Fig. 4. Recorder simplified schematic

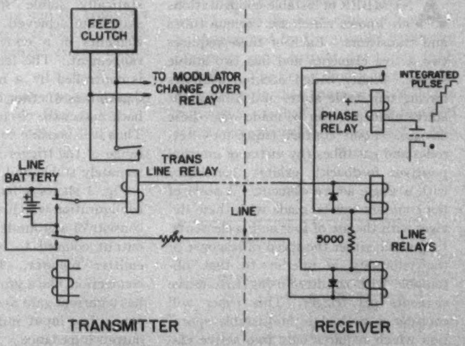


Fig. 5. Control circuit

pair of contacts of the phasing relay short-circuits the interrupting commutatator so that no further pole slippage can take place. A second pair of contacts on this same relay energizes the line feed clutch and recording of a properly framed message takes place.

It should be pointed out that in an area of 3-phase power distribution, a transmitter connected to phase 1 may have to send to a recorder connected to phase 2 or 3. Now it must be remembered that the angular positioning of the phasing slug is made in respect to a particular angle of current flow through the motor. This alignment, however, is done using one phase of alternating current, so a field adjustment is provided to index the angular position of the magnetic slug in steps of 120 degrees.

Control Circuit

Control-circuit operation, Fig. 5, is on a metallic loop basis because of the very heavy ground currents encountered along an electrified railroad way. A line battery of 120-volt d-c is used for calling and signalling and, for this reason, the primaries of all line-bridging transformers are split and a capacitor is inserted to block the d-c path. The line battery source is at the transmitter and it is connected to the line through the coil of a line relay. This transmitter line relay is adjusted to operate at 35 milliamperes.

Depressing either the message or coupon button of the transmitter serves to reverse the line battery polarity and thus bring up either the message or coupon line relay at the receiver. At the same time the message or coupon is wrapped around the transparent cylinder of the transmitter, and the exciter lamp and scanning motor are energized. As described previously, the modulator changeover relay is not energized, so only phasing pulses appear on the line. The line current at this time is limited to a value of less than 20 milliamperes by the line, a 5,000-ohm resistor in the receiver, and an adjustable resistor in the transmitter.

At the receiver, the operation of one of the line relays brought up either the message or coupon lamp and a buzzer indicating an incoming call. The receiving operator is thus advised to insert either a blank message form or a preprinted coupon form. Upon depressing the start button, the Teledeltos* recording blank will be automatically wrapped around the transparent cylinder, and power will be turned on to all receiving circuits. Phasing now takes place and a third pair of contacts on the phasing relay short-circuits the 5,000-ohm current-limiting resistor in the line.

The line current is thus increased to more than 40 milliamperes and operates the line relay of the transmitter. Battery is now disconnected from the line, the feed clutch is energized, and the modulator changeover relay operates to remove phasing pulses and transmit the message.

The 360-rpm transmitter uses a carrier frequency of 2,000 cycles and a frequency band extending from 560 to 3,440 cycles. Since oscillator and filter circuits of the modulator are on a plug-in basis, the 1,800- and 360-rpm transmitters are alike except for motor gearing. At 360 rpm the low velocity of the recording stylus permits the use of a mechanical clutch for phasing. The incoming phasing pulses are amplified and detected and directly used to trigger a relay tube which in turn energizes the phasing clutch.

Ticketfax machines are relatively new so that maintenance and operating logs for only several months are available. These records show only minor troubles, most of which can be attributed to personnel unfamiliarity with new equipment.

Reference

1. Ticketfax Transmitters and Recorders, D. M. Zabriskie. *AIEE Transactions*, vol. 74, pt. I, Sept. 1955, pp. 536–41.

---

* Registered trade-mark, Western Union Telegraph Company.

A Multistable Transistor Circuit

R. A. HENLE
NONMEMBER AIEE

A NUMBER of bistable configurations are known which use vacuum tubes and transistors. Each of these requires two active elements and has two stable states. Binary coded decimal counters having ten stable states and using eight active elements, can be made with these devices. Point contact transistors, tetrodes and gas tubes (by virtue of internal positive feedback) exhibit bistability with a single active element. A scale of ten counters can be made with these devices with the use of four active elements. In general, either the speed of response or the reliability is inferior to that obtainable with counters using two active elements per trigger. This paper will consider a method of multistable operation which requires only two active elements, theoretically, for any number of statically stable states. These stable states are achieved by using two active elements in a conventional trigger arrangement. The feedback in the circuit is controlled by a nonlinear load which determines whether the net positive feedback makes the circuit stable or unstable. Thus it is possible to divide the operating range of the trigger into a number of alternately stable and unstable regions.

Fig. 1 shows a basic transistor trigger configuration which will be analyzed. It consists of a grounded base stage with its output coupled to its input through an emitter follower. The grounded base connection, like a grounded grid amplifier, has a current gain approximately equal to one, a low input impedance, and a high output impedance. The emitter-follower stage, like a cathode follower, has a high input impedance, a low output impedance, and a voltage gain approximately equal to unity. Since neither the grounded base or the emitter-follower stage is inverting, positive feedback is obtained by returning a tap on the emitter follower output to the input of the grounded base stage.

A feedback factor AB can be defined, as in Fig. 1, where

$$A = \frac{i_{\text{out}}}{i_{\text{in}}} \qquad (1)$$

$$B = \frac{i_{\text{feedback}}}{i_{\text{out}}} \qquad (2)$$

if $AB < 1$ the circuit is stable, and if $AB > 1$ the circuit is unstable.

Paper 55-497, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 4, 1955; made available for printing June 21, 1955.

R. A. Henle is with the International Business Machines Corporation, Poughkeepsie, N. Y.

The writer wishes to acknowledge the many helpful suggestions of J. Logue while this work was in process.

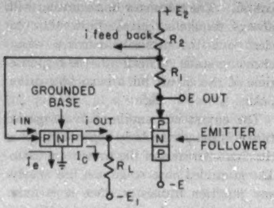


Fig. 1. Transistor trigger circuit

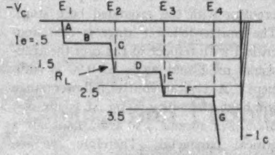


Fig. 2. Collector V-I characteristic with step load line

For small signals

$$i_{\text{out}}=dI_c \tag{3}$$

$$i_{\text{feedback}}=dI_{\text{feedback}} \tag{4}$$

For a junction transistor in the grounded base connection

$$\alpha=\frac{i_{\text{out}}}{i_{\text{in}}}=A \tag{5}$$

Assume that the emitter of the grounded base stage is at ground potential when biased in the forward direction. Then

$$I_{\text{feedback}}=\frac{E_{\text{out}}}{R_1}+\frac{E_2}{R_2} \tag{6}$$

$$E_{\text{out}}=I_cR_L-E_1 \tag{7}$$

Combining 6 and 7

$$I_{\text{feedback}}=\frac{I_cR_L}{R_1}-\frac{E_1}{R_1}+\frac{E_2}{R_2} \tag{8}$$

When equation 8 is differentiated, and equations 3 and 4 are used

$$\frac{dI_{\text{feedback}}}{dI_c}=R_L/R_1=\frac{i_{\text{feedback}}}{i_{\text{out}}}=B \tag{9}$$

Then

$$AB=\alpha R_L/R_1\approx R_L/R_1 \tag{10}$$

Therefore we see that for $R_L>R_1$ the circuit will be unstable and for $R_L<R_1$ the circuit will be stable. If then a load line could be constructed with alternate regions of $R_L>R_1$ and $R_L<R_1$, a circuit would result with alternate regions of unstable and stable operating points.

Consider a load line such as that in Fig. 2. If this load line is represented by $R_L=F(Ic)$, there are small values for R_L in regions *A*, *C*, *E*, and *G*, and large values for R_L in regions *B*, *D*, and *F*. Therefore the circuit can be stable in regions *A*, *C*, *E*, and *G*, and unstable in regions *B*, *D*, and *F* by proper choice of R_1.

A load line may be made which will closely approximate that shown in Fig. 2 in several ways which are shown in Figs. 3 and 4. The first method shown in Fig. 3 utilizes conventional diodes. Here

$$-E_1<-E_2<-E_3<-E_4$$

The most negative potential for the collector of T_1 is $-E_1$. As this potential diode *a* is conducting and diodes *b*, *d*, and *f* are reverse-biased, the transistor would be operating in region *A* of Fig. 2. As the collector of T_1 is brought more positive than $-E_1$, diode *a* becomes reverse-biased and the collector of T_1 sees the resistance R_3 returned to the voltage $-E$, this would correspond to region *B*, Fig. 2. When the collector voltage of T_1 becomes equal to the voltage $-E_2$ minus the voltage drop across diode *C* caused by the current generator $-E$ and R_3, diode *b* becomes conducting. For a small collector voltage region both diodes *b* and *c* are conducting and the transistor is operating in region *C* of Fig. 2. As the collector of T_1 is brought more positive than $-E_2$, diode *c* becomes reverse-biased and the collector of T_1 sees the resistance and current of two current generators ($-E$ and R_3) in parallel. This would be region *D*, Fig. 2. As the collector is brought still more positive diode *e* and *d* conduct giving region *E*, Fig. 2, and as the collector becomes more positive than $-E_3$ the collector voltage passes through region *F* where the collector of T_1 sees the resistance and current of three current generators ($-E$ and R_3) in parallel. Finally diode *f* conducts as the collector becomes more positive than $-E_4$ and T_1 is operating in region *G* of Fig. 2.

The second method, Fig. 4, utilizes silicon diodes. These diodes at a known reverse voltage (which may be controlled in manufacture) change from a very high impedance to a very low impedance. The voltage at which this change occurs is often referred to as the Zener voltage of the diode. The dynamic impedance of the diode in its Zener region is less than when the diode is forward-biased. By utilizing these diodes as shown in Fig. 4, advantage is taken of both the forward and Zener characteristic. The magnitude of the voltage step for the step load line is determined by the Zener voltage of the diodes. The magnitude of the current step is determined by $-E$ and R_3. Let R_f be equal to the forward resistance of the Zener diode and Rz be equal to the resistance of the Zener region. Then the impedance of the most negative voltage step for the circuit of Fig. 4 would be $Rz_1+Rz_2+Rz_3$. For the second step, diode 1 would be in its forward resistance region and diodes 2 and 3 in their Zener regions. The resistance of this step would then be $R_{f1}+Rz_2+Rz_3$. The third step would have a resistance equal to $R_{f1}+R_{f2}+Rz_3$, and the fourth step a resistance equal to $R_{f1}+R_{f2}+R_{f3}$. The high-resistance regions between steps would, as for the diode clamped load line, be determined by the parallel resistance of one or more current generators ($-E$ and R_3). The Zener diodes offer definite advantages in that the number of diodes required is one-half the number required when conventional diodes are used. Also

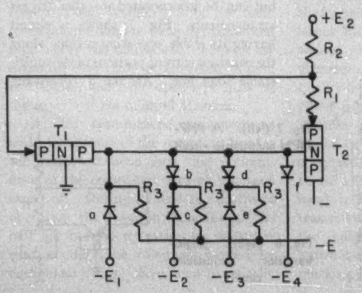


Fig. 3 (left). Diode clamped load

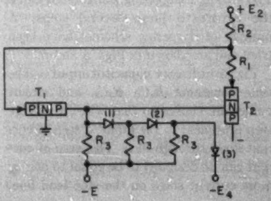


Fig. 4 (right). Zener diode load

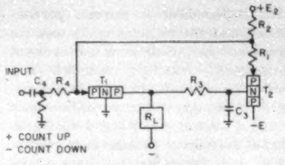


Fig. 5. Triggering method A

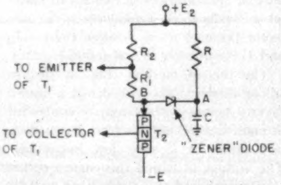


Fig. 6. Triggering method B

it is not necessary to have separate supplies or a low-impedance divider for supplying the step voltages. They have a disadvantage in that each low impedance region essentially has all the diodes in series, making it necessary to design the circuit with a larger value for R_1.

A typical design value for R_1 might be 1,500 ohms. For this case, using the diode clamped load line, equation 10 tells us that for stable regions to exist it would be necessary for the diodes used to have a dynamic forward resistance of less than 750 ohms. The parallel resistance of the current generators would have to have a value greater than 1,500 ohms. It is evident that these conditions do not present stringent requirements for either the diodes or the current generators.

Triggering Schemes

Precautions must be taken in the manner in which the circuits are triggered. Unlike a conventional trigger, which may be set or reset by pulses over a certain minimum amplitude, a circuit of this type obviously requires an upper limit on the energy of the input triggering pulse. Too strong a triggering pulse could cause the circuit to jump several steps. A number of triggering schemes have been tried and are shown in Figs. 5 and 6.

Fig. 5 utilizes a capacitor input. The time constants R_4C_4, R_3C_3, and circuit values are chosen so that when a voltage step is applied to the input, a step of current flows through R_L. This step of current can be designed to be equal to one or more current steps on the step load line. The circuit will count up for positive inputs and down for negative inputs. A disadvantage of this arrangement is that the rise time at the output is also limited by the time constant R_3C_3.

The triggering arrangement of Fig. 6 utilizes a Zener diode to quantize the trigger steps. The Zener voltage of the silicon diode is chosen to be equal to the step voltage. The quiescent voltage of point A is one step more positive than the quiescent voltage of point B. When a positive pulse is applied to the emitter or collector of T_1, the voltage at point B starts to rise, but is clamped by the Zener diode at the voltage of point A for a period sufficiently long to allow the circuit to stabilize at the new voltage. Condenser C will then charge to a voltage one step more positive than point B, and the circuit is ready for the next input pulse. A similar scheme could be arranged to make a count down input.

Another method of triggering would be to have the trigger input at the base of a PNP inverter stage which had its collector attached to the collector of T_1. Application of a negative pulse at the input would pull the collector of T_1 to the emitter potential of the inverter; this potential would be adjusted to be one step more positive than the collector potential of T_1. An NPN transistor could be used in a similar circuit for a count down input.

Speed of Operation

There appears to be no reason why circuits of this type cannot be as fast or faster than more conventional circuits using similar transistors. An advantage of the multistable circuit is that if properly designed neither transistor is permitted to go into saturation; thus, minority carrier storage problems are eliminated. In a conventional transistor trigger with two stable states (not employing special circuitry), one transistor will normally be in saturation and one transistor normally cut off. The transistor in saturation will always require a certain finite time to clear out the minority carriers when changing state. This limits the response time of the circuit for a particular drive pulse.

The circuit has another advantage in that it is not particularly dependent upon the characteristics of the transistors used. The grounded base stage can use nearly any junction transistor since it is only necessary that α be close to one. Transistors with $\alpha>0.95$ are more than adequate. Large values of I_{c0} and the temperature dependence of I_{c0} present no problem with the grounded base stage. The emitter-follower stage is also noncritical with respect to the α of the transistors used, and again $\alpha>0.95$ is sufficient. A typical emitter-follower stage will have a stability factor S^1, where $S=\partial Ic/\partial I_{c0}$ of about 1.3 so again I_{c0} is not a critical parameter. Therefore, we see that transistors with wide variations in α and noncritical specifications on I_{c0} are more than adequate. A circuit having ten stable states has been constructed and tested in the laboratory. This circuit used the step load line scheme of Fig. 3. The voltage steps were set at 1.5 volts and the current steps at 1 milliampere. Low-power audio-type transistors were used with f_{c0} of less than 10^6 cycles per second. With these transistors the circuit was operated at an input pulse rate of 60 kilopulses per second. A second ten-state counter was built with the use of a 5-stable-state circuit combined with a binary trigger to achieve 10 stable states. This circuit was so arranged that after the ninth input pulse a diode gate was conditioned which allowed the circuit to be reset by the tenth input pulse. Both of those circuits showed good interchangeability of transistors.

Other Multistable Configurations

The principle of multistable operation is not limited to the configuration shown but can be incorporated in other trigger arrangements. Fig. 7 shows a circuit having its stable operating regions along the constant current portions of the multistable load line. Assume T_1 operating

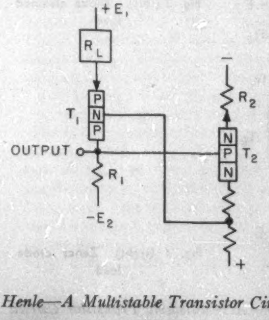


Fig. 7 (left). A PNP-NPN multistable circuit

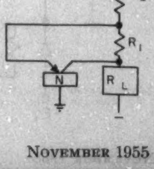


Fig. 8 (right). Single transistor multistable circuit along a constant current region. Then variations in the base voltage of T_1 will cause little or no variation in the output voltage and therefore no change in the voltage fed back to the base of T_1. Hence, the circuit is at a stable operating point. However, if T_1 were operating along a constant voltage region of the multistable load line, then small changes in the base voltage of T_1 would cause larger changes in the output voltage. These changes are fed into the base of T_2 causing changes in the collector voltage of T_2 which are fed back to the base of T_1. The circuit is unstable in this region.

Fig. 8 illustrates a single transistor trigger ($\alpha>1$) with a step load giving multistable operation. The operation of this circuit would be similar to that of the first circuit described. In general, conventional point contact transistors would not be satisfactory since a predictable uniform α and pentodelike characteristics are most desirable for the circuit.

Conclusion

This circuit appears to be well suited to transistor characteristics. The circuit utilizes those characteristics of junction transistors which are inherent in the device and do not require close control during manufacture. Desirable features of the circuit are that it can theoretically have any number of statically stable states and that these states may be achieved with the use of only two transistors with α less than 1 or one transistor with α greater than 1. This circuit provides an economical application of transistors to counter circuits.

Reference

1. Principles of Transistor Circuits (book), R. F. Shea. John Wiley & Sons, Inc., New York, N. Y., 1953, pp. 44–48.

Engineering Problems and Performance Record of the Hawaiian Interisland Radio Network

W. B. FARINON
NONMEMBER AIEE

A. C. WALKER
NONMEMBER AIEE

ol. 74
. 1
955

INSTALLATION of the present interisland radio network[1,2] of the Hawaiian Telephone Company required consideration of many equipment and transmission engineering problems. Because of the unique geographical configuration of the area served, a great deal of preliminary investigation was needed to select operating frequencies which were both available and suitable for use from a transmission standpoint. Various types of radio equipment were then considered to determine which could best provide the needed facilities. Selection of sites and installation of radio equipment also posed a number of specific problems. With the exception of 900-megacycle (mc) links terminated at the exchange buildings in Honolulu and Wailuku, all of the radio terminals are remote from the connected exchanges; the multiplexed channels are extended over cable at Honolulu, Kauai, and Lanai, and over open-wire lines on the islands of Hawaii and Molokai.

Detailed transmission measurements and examination of the operating logs for the various stations show that performance of the completed network has been excellent. During its first 18 months of operation, the 2-section Oahu-Hawaii circuit has been in operation for 99.94 per cent of the time. Outages caused by equipment failure have been practically nonexistent since most links are fully stand-by equipped. There have been absolutely no outages recorded on any links because of fading. On the two links to Hawaii, which have control and supervision equipment, there have been no outages caused by equipment failures. Noise measurements made on individual channels show performance to be equal to or better than would be obtained on comparable circuits operated on an overland route over wire line or cable carrier systems.

Propagation Problems

Extensive studies were made by the Hawaiian Telephone Company to select the frequency band most suitable for interisland radio links. The basic problem was to provide telephone and telegraph channels with commercially acceptable quality and reliability on overwater radio paths from 51 to 104 miles in length. Because of the high space loss on the longer links (plus need for maintaining adequate fade margin), it was necessary to use a frequency band for which equipment was available with relatively high transmitting powers. Also, because of the fading characteristics of overwater paths, it was necessary to operate on frequencies at which normal fades caused by reflection would be within acceptable limits. Consideration also had to be given to a peculiar characteristic of island weather in which the normal northeast trade winds are replaced by a strong warm front with successive days of hot, calm, humid weather. This condition causes extremely erratic propagation because of changes in atmospheric refraction. And, finally, sufficient bandwidth had to be available to permit proper frequency separation of the various radio links required. To help determine the highest practicable band of frequencies and to substantiate available data on propagation in the very-high-frequency and ultrahigh-frequency regions, a series of propagation tests was made between various islands at frequencies of 150, 460, and 1,900 mc.

The contrast between propagation conditions at 150 and 1,900 mc was demonstrated in tests made between Kole Kole peak on Mt. Haleakala, Maui, and the penthouse of the Alakea Street exchange building in Honolulu. A profile chart of the path (plotted on 4/3 earth radius) is shown in Fig. 1. Elevations of these two points are 10,025 feet and 90 feet respectively. Path length is slightly over 110 miles. The path includes a slight obstruction about 4 miles from the Honolulu receiving site. This obstruction was not considered detrimental to the tests

---

Paper 55-535, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 25, 1955, made available for printing May 6, 1955.

W. B. Farinon and A. C. Walker are with the Hawaiian Telephone Company, Honolulu, Hawaii.

The authors acknowledge the assistance of J. P. Thornton, Electronics Engineer of the Hawaiian Telephone Company, in the development of the interisland system. Others who contributed materially are S. Liner, R. Alden, and Y. C. Pang, all Electronics Engineers.

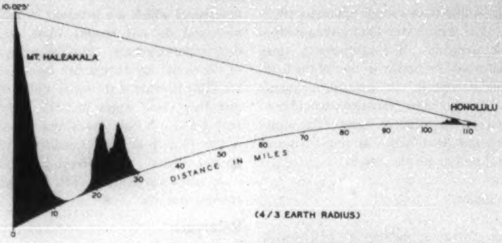


Fig. 1. Profile chart of the Honolulu-Haleakala path

since the path was selected for convenience rather than to provide optimum propagation. Calculations using the knife-edge diffraction theory indicated that the shielding hills would cause a loss of signal strength of about 6 decibels (db). With the equipment used for 1,900-mc tests, received signal strength at Honolulu should have been about 79 db below 1 watt; the average signal strength actually recorded was approximately 6 db higher than the calculated value. The difference could have been caused by any one of a number of factors, and no attempt was made to find the exact cause since the primary purpose of the test was to determine the fading range likely to be encountered in the various bands.

The 1,900-mc signal strengths (hourly variation) received over a 3-day test period are shown in Fig. 2. The lowest hourly variation was 6 db, and the highest was 27 db. Much of the time the signal fell below the threshold of the receiver used.

Contrasted with the received signal strength variations at 1,900 mc were the results of tests made over the same path using a frequency of 150 mc. Received signal strengths as recorded during one 4-day period is shown in Fig. 3. During the entire test, occasional variations of up to +7 db and −9 db from free space value were recorded. Over a much longer period, the amount of variation would probably be increased somewhat. However, because of the lower space loss at 150 mc and because higher transmitting powers could be used if necessary, it was determined that the lower frequency would be practicable for such a link but that 1,900-mc transmission would not be sufficiently reliable for commercial service. Several additional tests were made between various possible radio sites at 150 mc with, in all cases, results indicating that equipment operating at this frequency would provide satisfactory circuits.

A number of tests were also made between various sites at 460 mc, the most conclusive of which was between Kalepa, Kauai, and Mauna Kapu, Oahu. These sites had elevations of 2,650 and 400 feet respectively and provided a path length of 91 miles, as shown in Fig. 4. Results of the Kalepa-Mauna Kapu tests for a 6-day period are shown in Fig. 5. During the entire time in which signals were recorded, the range of fading was from about 5 db above calculated free space loss to about 32 db below. An interesting aspect of the 460-mc test which is apparent in Fig. 5 was the daily recurrence of a definite, almost sinusoidal, fading cycle. This was attributed to a condition of subnormal refraction caused by the sun which had the effect of apparently lowering the transmitter site. The same phenomenon appeared in other tests made at 460 mc.

Propagation tests made by the Hawaiian Telephone Company were undertaken to determine whether any or all of the frequency bands investigated were suitable for interisland commercial service rather than to determine the detailed propagation characteristics at the frequency concerned. In most cases, tests were conducted only long enough to determine whether the magnitude of fading was sufficient to interrupt transmission. Sufficient data were accumulated to determine that the use of 1,900 mc on the type and length of path involved was completely out of the question. Although 460-mc transmission might have been satisfactory for some of the contemplated links, this frequency band was also considered to be impractical since no equipment available at that time would have provided reliable transmission for all links with the signal-to-noise ratio required for commercial telephone service. However, the results obtained at 150 mc were sufficiently conclusive to indicate that, with reasonable care in the engineering of individual links, equipment operating in the very-high-frequency band could be used to provide acceptable circuits between all of the islands.

Equipment Considerations

To provide the circuit groups required over the various interisland circuits, broad-band radio equipment was obviously required. Four possible types of

Vol,
Pt. 1
1955

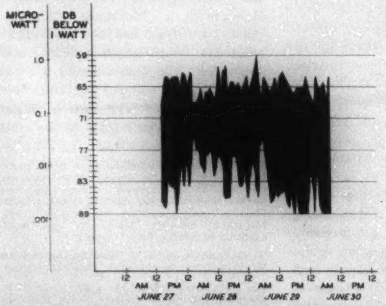


Fig. 2. Variation in received signal strength at 1,900 mc, Haleakala-Honolulu, June 27–29, 1950

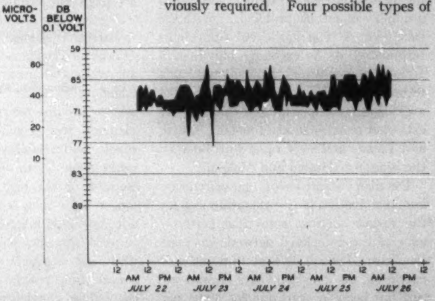


Fig. 3. Variation in received signal strength at 150 mc, Haleakala-Honolulu, July 22–26, 1950

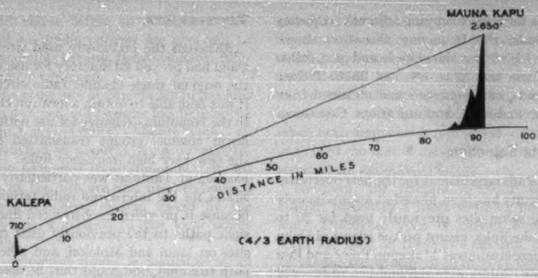


Fig. 4. Profile chart of the Mauna Kapu-Kalepa path

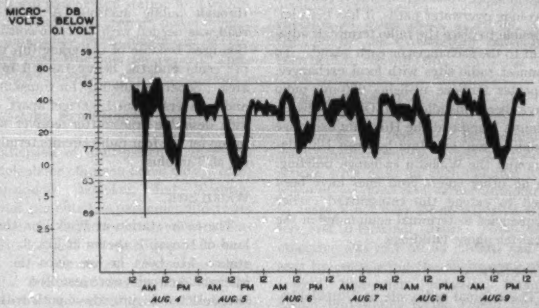


Fig. 5. Variation in received signal strength at 460 mc, Kalepa-Mauna Kapu, August 4–9, 1950

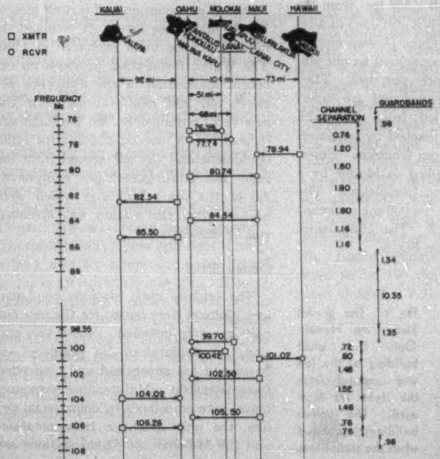

KAUAI
OAHU
MOLOKAI
MAUI
HAWAII
□ XMTR
○ RCVR
KALEPA
TANTALUS
HONOLULU
MAUNA KAPU
PUU
LANAI
LANAI CITY
PUU
92 mi
104 mi
73 mi
51 mi
68 mi
FREQUENCY Mc
76
78
80
82
84
86
88
98.35
100
102
104
106
108
76.98
77.74
78.94
80.74
82.54
84.54
85.50
99.70
100.42
101.02
102.50
104.02
105.50
106.26
CHANNEL SEPARATION
GUARDBANDS
.98
0.76
1.20
1.80
1.80
1.80
1.16
1.16
1.34
10.35
1.35
.72
.80
1.48
1.52
1.48
.76
.76
.98

Fig. 6 (left). Frequency allocation, interisland radio circuits

radio systems were considered in the early planning stages. These were pulse modulation, amplitude modulation, single side-band modulation, and frequency modulation.

Although pulse-modulation systems were available, and offered a technical possibility, such systems required very wide bandwidths which were not likely to be available at frequencies low enough to be free from fading interruptions on the long overwater paths. A further disadvantage of pulse-type modulation was the desire to use frequency division carrier equipment for land-line connections between exchanges and radio sites.

Investigation of the possibilities of using amplitude-modulated radio disclosed that, although narrow band (6-channel) systems had been manufactured, wide-band amplitude-modulated radio was not available commercially. Use of equipment capable of transmitting only six voice channels per circuit would have required excessively large numbers of transmitters and receivers at the Oahu and Maui radio sites. Consideration of single side-band equipment disclosed that, although such equipment was then available for the lower frequency bands, no development work had been carried out for single side-band radio capable of operation in the very-high-frequency band.

The only practical solution to the problem of furnishing the required interisland service proved to be wide-band frequency-modulated radio. Equipment using frequency-modulated techniques was available with bandwidth sufficient for transmitting 24 voice channels. Furthermore, use of frequency-modulated radio with frequency division carrier permitted more ready integration of the interisland radio network with the land-line toll circuits on the various islands. Frequency-modulated radio equipment capable of transmitting even more channels has also become available during the past few years,

Fig. 7 (right). A single building houses equipment for both overwater and overland links of the Oahu-Maui radio circuit

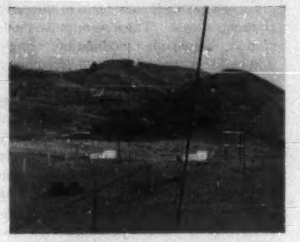


Fig. 8. Single building on left is used at Ulupalakua, Maui, for the repeater on the Oahu-Hawaii radio circuit. Cable in foreground connects to nearby Puu Mahoe repeater for Oahu-Maui circuit

with full use made of such new developments.

The contemplated interisland network included a minimum of 14 transmitters and 14 receivers for overwater links, all oriented approximately along the same line. Concentration of equipment at one Oahu site and at one Maui site was also considered desirable from an operating standpoint. Consequently, careful consideration had to be given to selection of operating frequencies which would provide suitable separation between channels and which would also provide a suitable guard band between the received and transmitted blocks of frequencies at each location.

It was determined that, with frequency-modulated radio equipment then available, suitable frequency separation between channels could be obtained by using one band of approximately 10 mc for transmitting from Oahu and Hawaii (receiving at Maui, Lanai, Molokai, and Kauai), and another band of approximately 10 mc for receiving at Oahu and Hawaii (transmitting from the other islands). It was further determined that interference between transmitters and receivers located at the same site could be prevented if the transmitting and receiving frequency bands were separated by a guard band of approximately 10 mc.

With these requirements in mind, plus the requirement for adequate guard bands between the interisland frequencies and the services occupying adjacent frequency bands, the frequency allocation shown in Fig. 6 was developed, and permission to use the 76 to 88- and 98 to 108-mc bands was requested and obtained from the Federal Communications Comission.

Site Selection

Radio equipment for the new overwater circuits has been placed at approximately the same sites previously used for 30 to 50-mc links, except on the island of Oahu where stations at Makapuu Point and Puu Manawahua were relinquished in favor of new locations at Mt. Tantalus and Mauna Kapu. In most cases, however, new buildings were constructed. Because line-of-sight transmission was required on the long overwater paths, it has been impossible to place the radio terminals adjacent to the exchanges on each island. To connect radio sites with local exchanges, repeater sections utilizing 900-mc radio have been installed on Oahu between Mauna Kapu and the Honolulu exchange building, and on Maui between Ulupalakua and the Wailuku exchange building. In all other cases, land lines have been used to extend the transmitted carrier frequencies to terminal equipment in the local telephone buildings.

MAUI SITE

The original radio site at Ulupalakua, Maui, did not provide a line-of-sight path to the Wailuku exchange. Consequently, equipment for the new Maui circuit has been installed in a new building on Puu Mahoe, a small knoll 1,500 feet from the original station which is still used as a repeater site for the Oahu-Hawaii circuit. Both locations are considered as one station for maintenance purposes. They are operated unattended and are supervised and controlled from Honolulu. The station can be reached by personnel in less than 1 hour from Wailuku. Emergency power is supplied to Mahoe by a connecting cable from Ulupalakua. The Mahoe building and the 900-mc antennas are shown in Fig. 7. The Ulupalakua buildings are shown in Fig. 8.

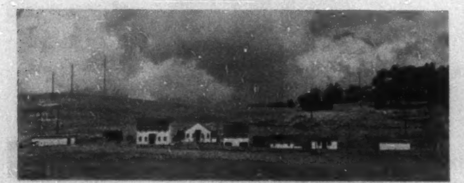


Fig. 9. The Waikii Station on Hawaii. Only the small building at the left and second from the right are now used. The other buildings, required when the station was attended, will be demolished

TANTALUS SITE

Although the previously used sites on Oahu had proved satisfactory for use with the original single-channel radio circuits, it was desirable to obtain a location closer to the Honolulu exchange for use with the larger channel groups transmitted over the new very-high-frequency links. The location at Tantalus was particularly desirable for the interisland radio terminals because it provided unobstructed line-of-sight paths to the previously used radio sites on Maui and Molokai, and a better path to Lanai, and could thus be used as the terminus for all circuits except those to Kauai.

The site at Tantalus, located in a Territorial Forest Reserve, was acquired through public auction. A 4,000-foot road was needed with concrete construction used because of the grade (up to 28 per cent) and the heavy rainfall in the area. Separate buildings for transmitters and receivers, about 1,500 feet apart, provide desirable transmitter-receiver separation for the four radio circuits terminating at Tantalus.

WAIKII SITE

The radio station at Waikii on the island of Hawaii is shown in Fig. 9. This station has been in use since the first interisland circuits were installed. The site, selected to provide a path without land intervention for the old nonrepeatered circuits to Oahu, now provides a line-of-sight path to Ulupalakua on Maui. Waikii is on the slopes of the 13,000-foot Mauna Kea mountain and is itself almost 1 mile above sea level. It is about 45 miles from the main exchange in Hilo. The Waikii installation provides on excellent illustration of the difference between old and new stations. For the new radio circuit only two buildings are in use, Fig. 9, one on the left for radio equipment and the one second from the right for power-generating equipment. All of the other buildings which were needed when the station was attended will be demolished.

OTHER SITES

The original radio stations on Lanai and Molokai were reused for the new circuits recently installed. These two stations are maintained by personnel on Maui and are supervised and controlled from Honolulu. Maintenance personnel travel between islands by commercial airline, the only available transportation, and the Molokai and Lanai stations are visited only monthly for routine maintenance. Although almost a full day

Vol
Pt.
195

might elapse before an emergency visit could be made, outages are expected to be negligible because the stations are fully stand-by equipped.

The radio site at Mauna Kapu has also been retained as a repeater location for the Oahu-Kauai circuits. Mauna Kapu can be reached in about 1 hour by maintenance personnel from Honolulu. Buildings at the site are also used to house a mobile telephone system leased by the telephone company to the local power utility.

Commercial power is available at all stations except Waikii, Hawaii, where the primary power source is a Diesel generator. Each station is equipped with a completely automatic emergency plant which will pick up full station load within 1 or 2 seconds after primary power failure. The emergency power plants can be controlled from Honolulu for periodic exercising and checking.

Land-Line Extensions

Extension of multiplexed channels to the telephone offices in Lanai City, Lanai, Kaunakakai, Molokai, and Waimea, Hawaii, presented no particular difficulties because the distances involved were relatively short and only small groups of channels were concerned. However, some special treatment was required for the open-wire lead extending from Waikii to Hilo on Hawaii and on the two cables to Tantalus on Oahu.

WAIKII-HILO EXTENSION

Six channels on the Oahu-Hawaii circuit are extended 13 miles to the Waimea exchange, and the remaining 12 channels are extended at carrier frequencies 45 miles to Hilo. Special separation filters for the two carrier groups are provided at Waikii. Point-type transposition was installed on the Waimea lead to permit operation on it of both east and west 3-channel *32E* systems.

The Waikii-Hilo extension is made on a 15-pair 104-mil copper voice-frequency transposed lead. This lead runs through a lava field for almost its entire length and is in an extremely wet area. The 12-channel type-*42C* carrier system utilizes frequencies of 36 to 84 kc in one direction and 100 to 148 kc in the other so no crosstalk problem exists. A special repeater with level regulation has been installed at Waikii to compensate for losses in the open-wire extension.

HONOLULU-TANTALUS EXTENSION

The transmitting and receiving cables extending between the Alakea exchange and Mt. Tantalus are used for four independent radio circuits. Identical channelizing equipment is used on the Lanai and Molokai circuits with the multiplexed channels occupying a frequency band of 4 to 24 kc in both directions. Since both frequencies and levels are coordinated, no special cable treatment was required for these two circuits.

However, the circuits to Maui and Hawaii presented a more complex problem. Different types of carrier equipment are used on these two circuits, one occupying frequencies of 36 to 84 kc transmitting and 100 to 148 kc receiving, and the other occupying frequencies of 40 to 88 kc in both directions. Except for addition of equalizers, no special cable treatment has been required. Through proper selection of cable pairs, crosstalk contributed by the transmitting cable has been held to below 60 db.

Link Engineering

Each of the overwater radio circuits was planned to provide normal toll quality service with noise in each channel contributed by radio and carrier equipment to average not over 29 db adjusted on a terminal-to-terminal basis. How this objective was met on all circuits can be seen by considering the engineering of a typical link—that between Kalepa, Kauai, and Mauna Kapu, Oahu, for example. Actually, two identical radio circuits operate over this route which has a path length of 92 miles.

The radio equipment installed was designed for operation at any frequency in the 75 to 110-mc region. It has a power output of 100 watts, with a minimum input signal strength of 250 microvolts required to maintain the per-channel noise below 29 db adjusted. Free space loss on the 92-mile path between Kalepa and Mauna Kapu was calculated to be 116 db. With 5-element Yagi antennas, a gain of 16 db for both ends of the circuit was assumed. *RG17/U* cable was used to connect equipment and antennas, with the total losses for the link of about 2 db. Radio equipment gain was calculated to be 109 db. Therefore calculations showed a 7-db fade margin (125—118) for 29 db adjusted circuits. Actual measurements with a *2B* noise-measuring set, *F1A* weighting, show readings of 14 to 23 db adjusted in both directions.

Reliability

Although the complete interisland network has not been in operation long enough to provide over-all performance and reliability statistics, a good approximation of the ultimate performance record to be expected can be obtained from those circuits which were placed in service during 1953 and early 1954. The first new circuits installed have now operated successfully for about 2 years.

Before the present very-high-frequency network was installed, lost circuit time on the interisland circuits was fairly high. On the three circuits previously existing to the island of Hawaii from Oahu, two of which operated at 30 to 50-mc and one at 152 to 162 mc, noise levels of 50 db adjusted and above were frequently experienced.

The repeatered radio circuit between Oahu and Hawaii is of particular interest because it operates over the longest paths (104 and 77 miles) and carries the heaviest traffic. During the first 18 months of operation this circuit was out of service only 424 minutes (99.94-per-cent reliability). Transmission predictions have been confirmed by the fact that absolutely no outages have been caused by fading of the received signal on any of the links involved.

About one-fourth of the total outage time on the circuit was caused by faulty operation of the transfer equipment. Most outages of this nature occurred during the first 6 months of operation and the cause has since been corrected.

Slightly more than one-fourth of the total outage time has been caused by power failure. Most of this was because of a faulty emergency generator at the unattended Ulupalakua repeater station. Human errors, such as misdialing on the control system, were also responsible for

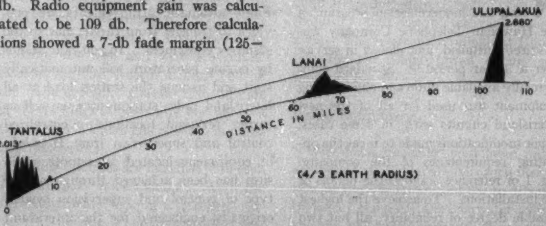


Fig. 10. Profile chart, Tantalus to Ulupalakua

ol. 74
. 1
55

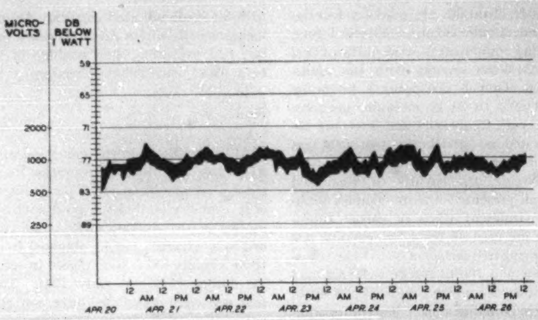


Fig. 11. Hourly fading chart for one 7-day period, April 20–26, 1953, Ulupalakua to Tantalus. Frequency 105.5 mc

about one-fourth of the total outage time. The remaining service interruptions have been due to a variety of causes.

Although the Maui, Molokai, and Lanai circuits have not yet been in operation long enough to provide comprehensive performance data, present indications are that they will be comparable to the Oahu-Hawaii circuit.

Since the two circuits between Oahu and Kauai are not equipped with stand-by, they cannot be directly compared with the other circuits of the interisland network; each circuit includes one 900-mc section and one overwater 76 to 108-mc section. During their first 6 months of operation, a total of 820 minutes of lost circuit time caused by equipment troubles was recorded on both Kauai circuit, with no extensive outages resulting from other causes. Since only one circuit would normally be out at any particular time, over-all service continuity between Oahu and Kauai is very good. If the total outage time is considered to be divided equally between the two circuits, each can be assumed to have been in service 99.05 per cent of the time during the 6-month period. It is anticipated that service will improve also as more operating experience is obtained.

Fading has had little or no effect on any of the interisland circuits. Under normal conditions, noise on the individual channels has been well below objectives and, even during the worst fades experienced, noise has not reached objectionable levels.

To provide a more accurate indication of fading range on the long overwater paths, measurements on the Oahu-Hawaii circuit were made over a period of several days on the 104-mile path between Tantalus and Ulupalakua. A profile chart of this path is shown in Fig. 10, and an hourly fading chart for a 7-day portion of the recorded period is shown in Fig. 11. During this time the maximum variation in received signal strength was about 8 db. Other measurements indicate that the maximum fade experienced during 18 months of operation has been about 12 db. This variation in received signal strength is not sufficient to affect either the talking or pulse-transmission channels on the radio circuit.

Conclusions

Successful operation of the overwater circuits in the interisland radio network has demonstrated the feasibility of using frequencies in the very-high-frequency band for long overwater routes carrying relatively large circuit groups. With complete freedom from interruptions caused by fading, the interisland radio circuits provide the Territory of Hawaii with communications as good as or better than could be obtained in any comparable all-land area. Excellent transmission quality, routine overwater dialing, and ample channels to meet traffic requirements help provide the Hawaiian Telephone Company with a completely integrated telephone plant.

References

1. The Interisland Radio Network of the Hawaiian Telephone Company, W. B. Farinon. *AIEE Transactions*, vol. 74, pt. I, Sept. 1955, pp. 529–36.

2. Equipment for Interisland Radio Circuits, A. C. Walker, V. P. Hoover, E. E. Nolan. *Ibid.*, Nov. 1955, pp. 576–83.

Equipment for Interisland Radio Circuits

A. C. WALKER
NONMEMBER AIEE

E. E. NOLAN
NONMEMBER AIEE

V. P. HOOVER
NONMEMBER AIEE

NEW equipment for the 4-step expansion and modernization program of the Hawaiian Telephone Company was procured, installed, and placed in service over a 3-year period.[1,2] Standard commercially available radio and channelizing equipment was used for all of the new interisland circuits with, in some cases, minor modifications made to meet the operating requirements of the company. Fig. 1 of reference 1 shows the details of the installations. To achieve the highest possible degree of reliability, all but two individual radio links are equipped with stand-by radio transmitters and receivers. Fully automatic arrangements are provided to transfer from the regular to the stand-by equipment. In addition, stand-by engine generators will automatically start and assume the station load at all interisland radio station sites as well as carrier terminal locations. Centralized control and supervision from Honolulu of equipment located at remote radio sites has been achieved through a new type of control and supervision system originally engineered for the interisland network. This system is of particular interest because of its flexibility in application and because of the manner in which it provides for control and supervision of a large number of functions with relatively simple equipment and circuit arrangements, and with a minimum number of transmitted tones. Control and supervision systems are already in service on the Maui and Hawaii circuits. Addi-

---

Paper 55-536, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 25, 1955; made available for printing May 16, 1955.

A. C. Walker and V. P. Hoover are with the Hawaiian Telephone Company, Honolulu, Hawaii; and E. E. Nolan is with the Lenkurt Electric Company, San Carlos, Calif.

The authors acknowledge the contributions of E. C. Schoen, Transmission Maintenance Engineer of the Hawaiian Telephone Company, to the development of the remote control equipment.

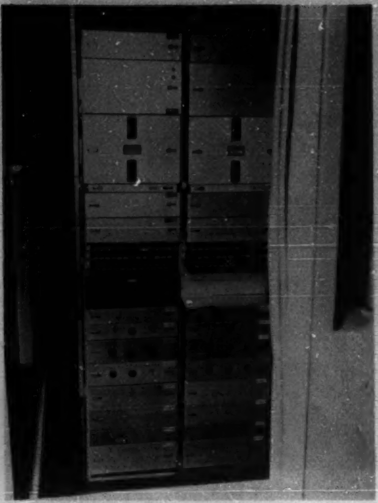


Fig. 1 (left). Type-42-6 carrier equipment at Honolulu

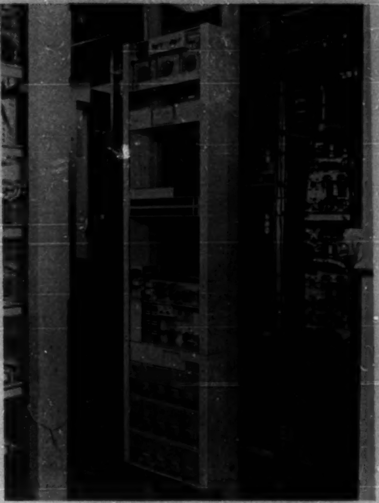


Fig. 2 (right). Type - 45BX terminal

l, 74
1
55

tional equipment is presently being installed to replace existing supervisory control equipment on the Lanai and Molokai circuits. Future plans include equipping the two circuits to Kauai with a control and supervision system to replace equipment now in use.

Channelizing Equipment

The number of 3-kc channels on the various interisland radio circuits varies from four on the circuit between Lanai and Oahu to 18 on the circuit between Oahu and Hawaii. At least one channel on each radio circuit is subdivided by means of carrier telegraph equipment for use with teleprinters. In all cases, the radio links have been multiplexed with standard frequency division carrier telephone equipment using single side-band, suppressed-carrier techniques.

In so far as possible, carrier equipment originally installed prior to 1950 has been reused on the new radio circuits. To achieve the most economical equipment arrangements, much of this older equipment which was first installed on the heavy traffic routes has now been shifted to the lighter traffic routes. For example, the Lenkurt *42-6* carrier telephone systems of the type shown in Fig. 1, originally installed on the old 152- to 162-megacycle (mc) circuits to Hawaii and Maui, are presently in service on the new circuits recently installed on Molokai and Lanai. Type *42-6* carrier equipment is also used on the two Kauai circuits.

During the first step of the expansion program, two types of wire line carrier systems (Lenkurt types *42C* and *32E*) were modified for use on the new radio circuit connecting Oahu and Hawaii. Wire line rather than radio carrier was used to simplify the problems involved in extending the multiplexed channels from the Waikii radio terminals to the Hilo and Waimea exchanges on the island of Hawaii. One *42C* system provides 12 voice channels to Hilo and two *32E* systems provide six voice channels to Waimea. The *32E* systems utilize frequencies between 7.3 and 17.0 kc in the east-west direction and 19.7 and 30.2 in the west-east direction, while the *42C* system operates between 36 and 84 kc in the west-east direction and 100 to 148 kc in the east-west direction. Adequate modulation bandwidth was available in the radio equipment used on this circuit, so the standard 2-wire allocation of the *42C* system was retained. Six circuits were obtained between 7.3 and 30.2 kc by utilizing east-west and west-east allocations of the *32E* system at both ends of the radio system. Careful transposition of the 13 miles of open-wire line between Waikii and Waimea has resulted in crosstalkfree circuits.

Dialing facilities are included with each *42C* and *32E* voice channel. With the *32E* equipment, form of interrupted carrier transmission is employed with operation on an amplitude-modulated basis. A frequency shift arrangement operating on a speech-plus basis is used with the *42C* channels.

The *42C* carrier will soon be removed from the radio link and reused as a wireline system to provide additional trunk circuits needed for extended area service on the island of Hawaii. It is anticipated that a Lenkurt *45BX* carrier system, capable of ultimate expansion to 48 channels, will replace the older equipment on the radio circuit.

A new 12-channel *45BX* radio carrier system has been installed on the Oahu-Maui radio circuit. One terminal of this carrier system is shown in Fig. 2. Dialing facilities between Honolulu and Wailuku terminals of the circuit are provided by built-in frequency shift signaling circuits associated with the individual carrier voice channels on an out-of-band, in-channel basis.

Most of the interisland teleprinter circuits are provided by Lenkurt type-*22A* frequency shift carrier telegraph equipment. The telegraph terminals at Honolulu for the Oahu-Hawaii circuit are shown in Fig. 3.

Radio Equipment

Wide-band frequency-modulated radio equipment is used on all links with operation in the 76- to 88- and 98- to 108-mc frequency ranges on the overwater paths, and the 890- to 940-mc range on overland paths. Equipment for both frequency

Fig. 3 (left). Type-22 telegraph carrier equipment at Honolulu

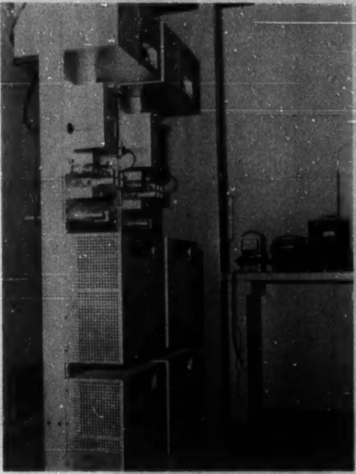


Fig. 4 (right). Radio equipment at Mauna Kapu

ranges is very similar in design and functional operation, with the primary differences being in the number of frequency multipliers provided and in the power output stages. Several versions of the same basic equipment design have been used, with full advantage taken of new developments made as the interisland expansion and modernization program has progressed. The equipment now in use from Oahu to Kauai, Lanai, and Molokai provides a 150-kc modulation bandwidth with a rated deviation of 150 kc and a rated power output of 100 watts for the 80-mc transmitters and 5 watts for the 900-mc transmitters. Equipment on the circuits from Oahu to Maui and Hawaii has a 300-kc modulation bandwidth with all other characteristics remaining essentially the same. Service channels associated with the radio equipment are used on all links. The radio equipment shown in Fig. 4 is typical of that used on all links.

To assure the highest practicable degree of reliability under the circumstances outlined in companion papers,[1,2] all of the radio links are fully equipped with stand-by. To avoid time-consuming switching between regular and spare equipment with consequent interruption of teleprinter circuits and dialing channels, individual antennas are provided for each transmitter and receiver. Five element Yagi antennas are normally mounted on 35 to 50-foot telephone poles using double crossarms approximately 10 feet long. Horizontal polarization is used. The antenna gain realized is approximately 8 decibels. Pole mounting of antennas has proved more suitable than pipe-mast mounting. For most locations, poles are more sturdy and much better antenna accessibility.

Transmitters are monitored by a micromatch unit inserted in the coaxial output line at the transmitter equipment rack. Transfer is initiated when the output power drops to 50 per cent of normal value with full transfer completed in as little as 75 milliseconds. A transfer initiated manually, either at the site or from the Honolulu control center, is completed in 20 to 40 milliseconds.

Receiver operation is monitored by special pilots transmitted over each link with a pilot signal receiver located at the output of each broad-band receiver. Loss of the pilot signal initiates transfer to the alternate receiver. Receiver transfer is completed practically instantaneously. Various safeguards are included in the transfer equipment to prevent transfer to an inoperative receiver and to prevent hunting between equipments in the event that both regular and stand-by units are inoperative. Receiver transfer normally has no effect on either voice or telegraph circuits. Transmitter transfer, however, might cause hits on the teleprinters or interrupt dialing in process. Receiver transfer will nor normally be initiated by a transmitter failure since Codan relays will release simultaneously in both the regular and alternate receivers upon failure of the distant transmitter. This keeps the receivers from switching unless there is a marked difference in release time. The transfer and monitor equipment arrangements for both transmitters and receivers are described in more detail in the following.

Remote Control and Supervision

With any network of radio circuits, it is desirable to have some system which will provide for centralized control and supervision of all stations in the network. Such a system not only provides more positive control over radio operation but it also permits the radio stations to be operated unattended while complying fully with the Federal Communications Commission regulations pertaining to radio operation. In the interisland radio network, a supervision and control system was particularly desirable since it would permit unattended operation of radio stations in relatively isolated locations. Major economies in operating costs could thus be obtained through concentration of personnel at the regular telephone exchanges. With the previous radio network, operating personnel had to be on duty constantly at most stations and, at many locations, housing and other facilities had to be provided by the company.

Any control and supervision system used had to overcome certain basic problems inherent in the interisland radio network. For example, since alternate routes were not available between islands, the control and supervision information

Fig. 5 (left). Type-51B equipment at Honolulu

Fig. 6 (above). Radio equipment for Tantalus and Ulupalakua and control and supervision equipment for all locations of the Oahu-Maui circuit

L 74
1
55

had to be transmitted over the same circuit being supervised. Also, arrangements were required which would permit both outgoing and incoming radio links to be controlled and supervised from one location. For complete control of the circuit, positive indication of link operation or failure was also required.

When investigation disclosed that no previously designed systems could fulfill all the requirements of the interisland network, the Hawaiian Telephone Company, in co-operation with the Lenkurt Electric Company, undertook the design of a new arrangement for transmission of supervisory and control information. The system devised, and now in use on the circuits from Oahu to Maui and Hawaii, has been standardized and designated by Lenkurt as the type-*51B* control and supervision system.

Each *51B* system can be used to control up to 20 operations and to supervise up to 20 alarm conditions at each of four remote stations from a single control center. With Honolulu as the control center, one of the systems now installed provides for control and supervision of radio stations at Tantalus on Oahu, Ulupalakua on Maui, and Waikii on Hawaii. The second *51B* system is associated with the radio circuit to Maui; remote stations are at Tantalus, Ulupalakua, and Wailuku. Since each of these systems is capable of providing control and supervision for one additional remote station, equipment is now being added for control of the Molokai radio circuit on one system, and for control of the Lanai radio circuit on the other. A third system will be added in the future with remote stations at Mauna Kapu and Kalepa, the radio sites for the two circuits between Oahu and Kauai. Supervision and control equipment at the Honolulu control center is shown in Fig. 5. Control and supervision equipment for Tantalus, Ulupalakua, Wailuku and Honolulu, and all radio equipment for the Tantalus-Ulupalakua link are shown in Fig. 6. The control and supervision sections of the *51B* system are essentially separate although the circuit arrangements are such that initiation of any remote operation through the control section will cause the supervision section to operate and record the change in conditions at the Honolulu control center.

Control System Operation

Basic elements of the control section of a *51B* system are shown in the simplified block diagram of Fig. 7. A telephone dial and signal transmitter are required at Honolulu, with a signal receiver and step switch at each remote location. A dial control is also added at each remote station to permit local control when desired.

Control is accomplished by dialing of a predetermined 2-digit code number which is then transmitted to all of the remote locations as a series of tone pulses. Only a single tone is needed for control of all stations in the system. Each station is assigned one or more groups of ten 2-digit numbers, with each number corresponding to some desired control operation. When the proper code number is dialed and transmitted through the system, a momentary ground (approximately 100 milliseconds) will be applied through the step switch to connected external equipment.

Supervision System Operation

Basic elements of the supervision section of a *51B* system are shown in the simplified block diagram of Fig. 8. For supervision, one tone is required to transmit information from the control center to all remote locations. One additional tone per remote location is needed to transmit information to the control center.

A change in condition of any supervised

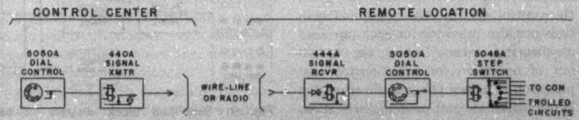


Fig. 7. Simplified block diagram of control section of 51B system

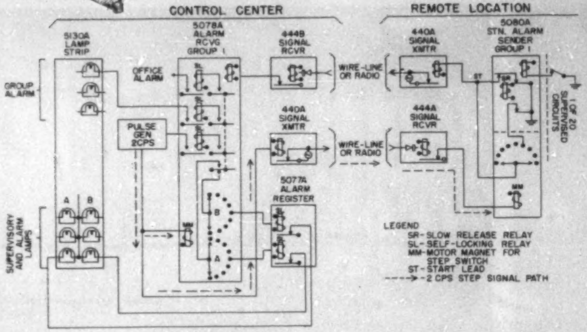

CONTROL CENTER
REMOTE LOCATION
5130A LAMP STRIP
5078A ALARM RCVG GROUP I
444B SIGNAL RCVR
440A SIGNAL XMTR
5080A STN. ALARM SENDER GROUP I
GROUP ALARM
OFFICE ALARM
SL
WIRE-LINE OR RADIO
ST
SR
1 OF 20 SUPERVISED CIRCUITS
PULSE GEN 2CPS
440A SIGNAL XMTR
444A SIGNAL RCVR
5077A ALARM REGISTER
MM
A
B
SUPERVISORY AND ALARM LAMPS
LEGEND
SR-SLOW RELEASE RELAY
SL-SELF-LOCKING RELAY
MM-MOTOR MAGNET FOR STEP SWITCH
ST-START LEAD
- - - -> -2 CPS STEP SIGNAL PATH

Fig. 8. Simplified block diagram of supervision section of 51B system

function will cause ground to be applied to a station relay and to one level of a rotary step switch on the station alarm-send panel at the remote location. This serves to interrupt the tone from the remote signal transmitter. Loss of tone at the control center will cause a group alarm to be actuated and, at the same time, start a pulse generator. Pulses, at a rate of two per second, are applied to both the step switch on the alarm-receiving panel and to the signal transmitter at the control center. Receipt of these pulses at the remote location causes the step switch on the station alarm-send panel to step in synchronism with the switch on the control center alarm-receiving panel. Indications of ground or no ground at the remote location key the signal transmitter for transmission of no tone or tone to the control center. This information is received by appropriate relays on the station alarm-receiving panel which, in turn, cause pulses to be applied to the proper levels of the control center step switch. From the step switch, information is passed to the alarm-register panel which controls the lighting of *A* and *B* lamps on the lamp display. Any change in conditions causes the affected lamps to flash, whereas lamps associated with unchanged conditions remain in their previous steadily lighted state.

The scanning cycle just described, which is automatically initiated upon any change in status of a supervised condition, can also be initiated manually from the control center. Other manual controls are also provided for such purposes as silencing alarms, locking the system out of operation for maintenance purposes, or to reduce power drain when the lamp banks are not under observation. Even when the system is locked out of normal operation, a change in a supervised condition will actuate a group alarm lamp plus any connected external office alarms.

In addition to the lamp bank display on the *51B* equipment rack, many of the lamp connections are paralleled to a graphic display of the entire radio network. This "mimic board" in the operating supervisor's office is also readily visible from the equipment floor. It shows at a glance the operational status of all equipment in the interisland network.

Application to Maui Radio Circuit

The *51B* control and supervision system on the Oahu-Maui radio circuit is typical of the systems already installed or proposed for installation on the interisland network. This radio circuit provides 12 voice channels terminated at the Alakea exchange in Honolulu and at the Wailuku exchange on Maui. The Honolulu carrier terminal and the control center for the *51B* system are located in the Honolulu main office building with all information transmitted to and received from the Tantalus radio site over two connecting cables. Thus, Tantalus becomes the first remote site associated with the system. An 80-mc radio link is installed between Tantalus and Ulupalakua, the second remote location. At Ulupalakua, the 80-mc terminal is connected back to back with a 900-mc terminal for a second link to Wailuku, the third remote location. Through the *51B* system, operation of equipment at all three remote locations is controlled, and conditions of the equipment are supervised. Special link-alarm arrangements for each of the four radio links are incorporated to back up the functions normally provided by the *51B* system.

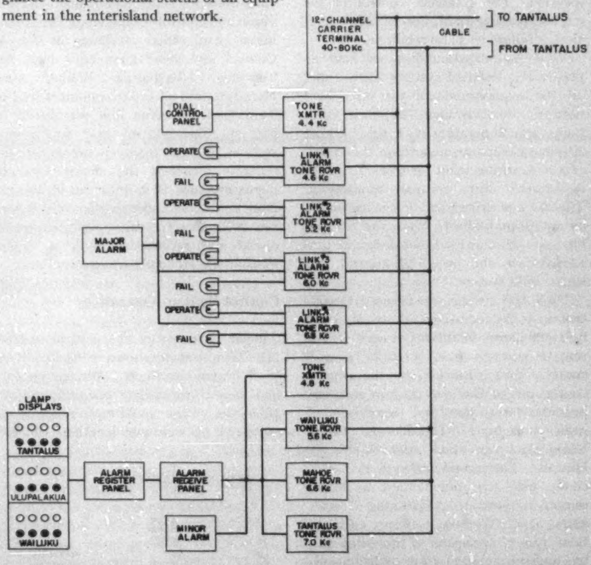

12-CHANNEL CARRIER TERMINAL 40-80 Kc
CABLE
TO TANTALUS
FROM TANTALUS
DIAL CONTROL PANEL
TONE XMTR 4.4 Kc
OPERATE
FAIL
LINK #1 ALARM TONE RCVR 4.6 Kc
LINK #2 ALARM TONE RCVR 5.2 Kc
LINK #3 ALARM TONE RCVR 6.0 Kc
LINK #4 ALARM TONE RCVR 6.8 Kc
MAJOR ALARM
TONE XMTR 4.8 Kc
WAILUKU TONE RCVR 5.6 Kc
MAHOE TONE RCVR 6.6 Kc
TANTALUS TONE RCVR 7.0 Kc
LAMP DISPLAYS
TANTALUS
ULUPALAKUA
WAILUKU
ALARM REGISTER PANEL
ALARM RECEIVE PANEL
MINOR ALARM

Fig. 9. Simplified block diagram of control, supervision, and transfer equipment at Honolulu

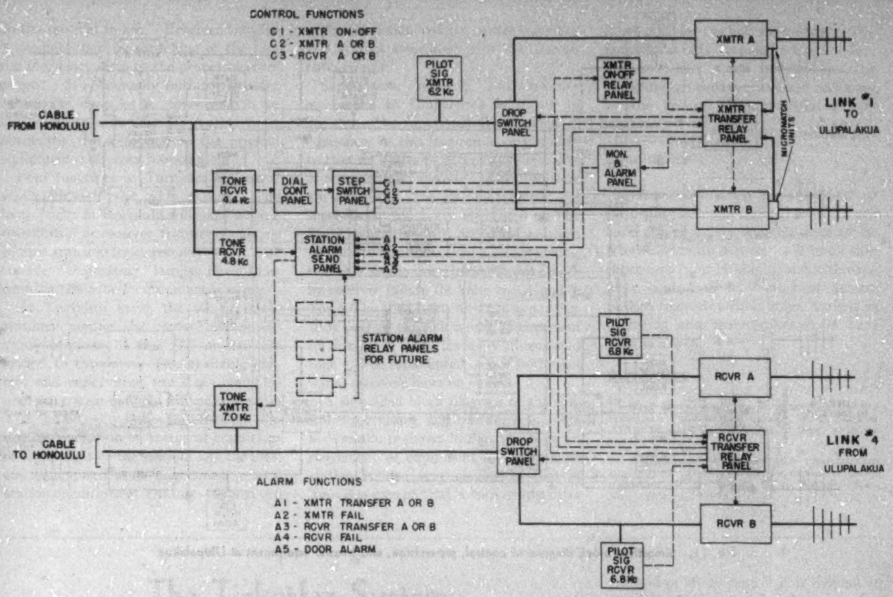

CONTROL FUNCTIONS
C1 - XMTR ON-OFF
C2 - XMTR A OR B
C3 - RCVR A OR B
PILOT SIG XMTR 6.2 Kc
CABLE FROM HONOLULU
DROP SWITCH PANEL
XMTR ON-OFF RELAY PANEL
XMTR A
XMTR TRANSFER RELAY PANEL
MICROWATCH UNITS
LINK #1 TO ULUPALAKUA
MON. & ALARM PANEL
TONE RCVR 4.4 Kc
DIAL CONT. PANEL
STEP SWITCH PANEL
C1
C2
C3
XMTR B
TONE RCVR 4.8 Kc
STATION ALARM SEND PANEL
A1
A2
A3
A4
A5
STATION ALARM RELAY PANELS FOR FUTURE
PILOT SIG RCVR 6.8 Kc
RCVR A
TONE XMTR 7.0 Kc
CABLE TO HONOLULU
DROP SWITCH PANEL
RCVR TRANSFER RELAY PANEL
LINK #4 FROM ULUPALAKUA
RCVR B
PILOT SIG RCVR 6.8 Kc
ALARM FUNCTIONS
A1 - XMTR TRANSFER A OR B
A2 - XMTR FAIL
A3 - RCVR TRANSFER A OR B
A4 - RCVR FAIL
A5 - DOOR ALARM

Fig. 10. Simplified block diagram of control, supervision, and transfer equipment at Tantalus

1, 74
1
55

A simplified block diagram of the control, supervision, and link-alarm arrangements at Honolulu is shown in Fig. 9. In addition to the dial control, alarm-receiving panels, and lamp banks, equipment at Honolulu includes four tone receivers for the individual link alarms, three tone receivers for supervisory information from the three remote locations, and two tone transmitters, one for control information, and one for supervisory information.

Only five different tones are used for the basic supervision and control functions of the *51B* system. Four tones are required for the link-alarm arrangements provided, with two additional tones used in the radio system as receiver monitoring pilots (pilot tones on the links incoming to Tantalus are also used to actuate the link alarms). The 11 tones required are spaced at 200- or 400-cycle intervals between 4.4 and 7.0 kc, thus falling within within the base-band limits of the wide-band frequency-modulated radio equipment. The service channel associated with the circuit occupies frequencies between about 250 and 3,000 cycles.

The link-alarm tones are not directly associated with the *51B* system although any condition which will cause a link alarm will also be registered through the *51B* supervision equipment. Loss of link no. 1, Tantalus to Ulupalakua, or link no. 2, Ulupalakua to Wailuku, will interrupt normally transmitting link-alarm tone transmitters to Ulupalakua or Wailuku. Loss of link no. 3, Wailuku to Ulupalakua, or link no. 4, Ulupalakua to Tantalus, will interrupt the pilot signals which are also used on those links for receiver monitoring. Interruption of any one of the four link-alarm tones will cause a major alarm to be given at the Honolulu office, and the specific link in trouble can be determined from observation of the operate and fail lamps. A system of interlocked relays is provided with the link-alarm tone receivers giving a positive indication of which of the four links has failed. Link alarms operate only when both regular and stand-by radio equipment is inoperative because of power failure or simultaneous failure of equipments. The supervision section of the *51B* system is arranged so that a minor alarm will be given whenever a change in remote conditions causes the automatic scanning cycle to be initiated.

A simplified block diagram of the control, supervision, and transfer arrangements at Tantalus is shown in Fig. 10. The transmitters for link no. 1 and the receivers for link no. 4 are at this site.

Control information originating in the Honolulu main office is received at Tantalus over the connecting cable as tone pulses which are changed to direct current and directed to the proper equipment through a step switch panel. Through the control arrangements provided, it is possible to transfer between the *A* and *B* transmitters and receivers to energize or de-energize the transmitters (the stand-by receiver is kept in continuous operation). Transmitter transfer is initiated by application of a ground which is dialed through the afore-mentioned step switch panel into the transmitter transfer relay panel. Transfer is accomplished by the application of cathode ground to one transmitter, removing cathode ground from the other, and switching the carrier frequency baseband input to the proper transmitter. Transmitter transfer will also be initiated automatically when the monitoring arrangement indicates that the transmitter in use is radiating less than 50 per cent of the normal power output.

Receiver transfer is initiated by the receiver transfer relay panel, either through the control system or by loss of the pilot

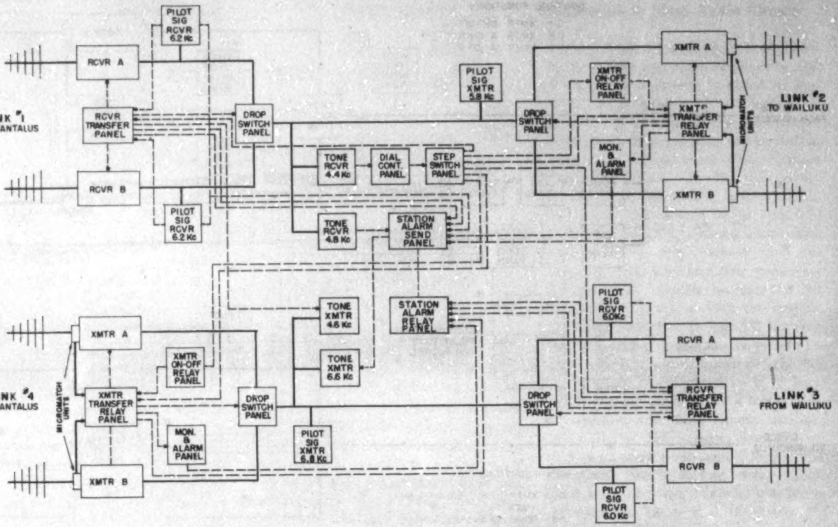

PILOT SIG RCVR 6.2 Kc
RCVR A
LINK #1 FROM TANTALUS
RCVR TRANSFER PANEL
RCVR B
PILOT SIG RCVR 6.2 Kc
DROP SWITCH PANEL
TONE RCVR 4.4 Kc
DIAL CONT. PANEL
STEP SWITCH PANEL
TONE RCVR 4.8 Kc
STATION ALARM SEND PANEL
PILOT SIG XMTR 5.8 Kc
DROP SWITCH PANEL
XMTR ON-OFF RELAY PANEL
MON. & ALARM PANEL
XMTR A
XMTR TRANSFER RELAY PANEL
XMTR B
MICROMATCH UNITS
LINK #2 TO WAILUKU
TONE XMTR 4.6 Kc
TONE XMTR 6.6 Kc
STATION ALARM RELAY PANEL
PILOT SIG RCVR 6.0 Kc
RCVR A
RCVR TRANSFER RELAY PANEL
LINK #3 FROM WAILUKU
DROP SWITCH PANEL
RCVR B
PILOT SIG RCVR 6.0 Kc
XMTR A
LINK #4 TO TANTALUS
MICROMATCH UNITS
XMTR ON-OFF RELAY PANEL
XMTR TRANSFER RELAY PANEL
DROP SWITCH PANEL
MON. & ALARM PANEL
PILOT SIG XMTR 6.8 Kc
XMTR B

Fig. 11. Simplified block diagram of control, supervision, and transfer equipment at Ulupalakua

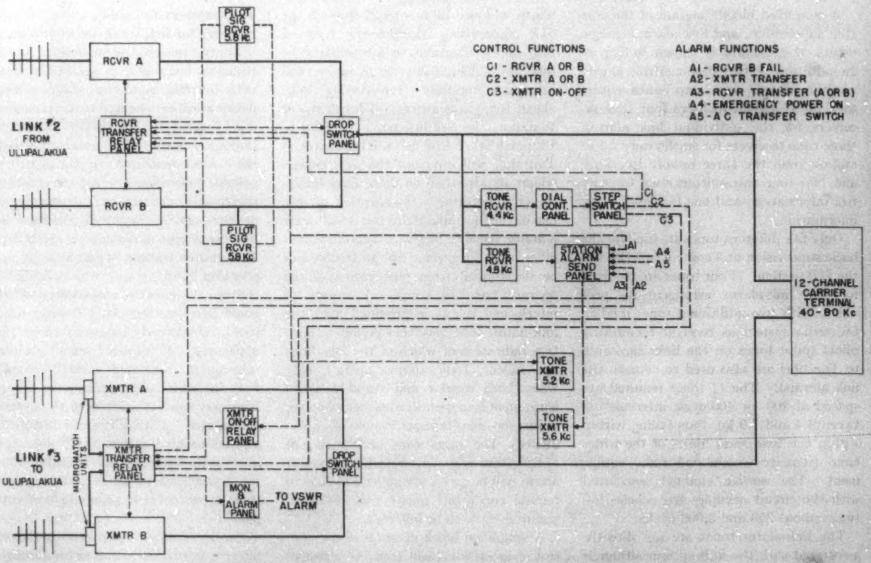

PILOT SIG RCVR 5.8 Kc
RCVR A
LINK #2 FROM ULUPALAKUA
RCVR TRANSFER RELAY PANEL
RCVR B
PILOT SIG RCVR 5.8 Kc
DROP SWITCH PANEL
CONTROL FUNCTIONS
C1 - RCVR A OR B
C2 - XMTR A OR B
C3 - XMTR ON-OFF
ALARM FUNCTIONS
A1 - RCVR B FAIL
A2 - XMTR TRANSFER
A3 - RCVR TRANSFER (A OR B)
A4 - EMERGENCY POWER ON
A5 - A C TRANSFER SWITCH
TONE RCVR 4.4 Kc
DIAL CONT. PANEL
STEP SWITCH PANEL
C1
C2
C3
TONE RCVR 4.8 Kc
STATION ALARM SEND PANEL
A1
A4
A5
A3
A2
12-CHANNEL CARRIER TERMINAL 40-80 Kc
TONE XMTR 5.2 Kc
TONE XMTR 5.6 Kc
XMTR A
LINK #3 TO ULUPALAKUA
MICROMATCH UNITS
XMTR TRANSFER RELAY PANEL
XMTR ON-OFF RELAY PANEL
DROP SWITCH PANEL
MON. & ALARM PANEL
TO VSWR ALARM
XMTR B

Fig. 12. Simplified block diagram of control, supervision, and transfer equipment at Wailuku

on the receiver in use. Receiver transfer is accomplished by switching of the carrier frequency drop to the proper receiver output. If a manually initiated transfer causes the drop to be connected to an inoperative transmitter (or receiver), an automatic transfer back to the original equipment will occur immediately.

Four functions at Tantalus are supervised through the *51B* system. The lamp banks at Honolulu will show when a transmitter or receiver transfer occurs, or when a transmitter or receiver fails. The transfer indicating lamps also show whether the *A* or *B* equipment is in use.

At Tantalus, as at the other remote stations, neither the control nor supervision sections of the *51B* system are loaded to capacity. The available control and supervision functions could be used for transmission of any other desired *A-B*, on-off, or yes-no information concerning operation or status of conditions or equipment at the station. At the present time, information concerning general station conditions at Tantalus is conveyed to the Honolulu control center over the *51B* system associated with the Hawaii radio circuit.

Supervision, control, and transfer equipment at Ulupalakua is shown in Fig. 11. The functional arrangement of equipment at this location is very similar to that at Tantalus except that two sets of receivers and two sets of transmitters are involved. The various control and supervisory functions are listed in the figure, with the only difference between Ulupalakua and Tantalus being the addition of a supervisory indication of standby receiver failure for links nos. 1 and 3. The 4.6-kc tone transmitter shown in Fig. 10 is part of the link alarm arrangement for link no. 1. The normally transmitted tone will be interrupted only if both pilot signal receivers have no output.

A simplified block diagram of the control, supervision, and transfer equipment at Wailuku is shown in Fig. 12. The arrangement of equipment at this remote station is also very similar to that at Tantalus except that a tone transmitter is provided for the link no. 2 alarm. The exchange is only attended 8 hours a day, and a 24-hour alarm indication at the Honolulu terminal is considered necessary by the Federal Communications Commission.

Conclusions

The control and supervision systems on the Oahu-Maui and Oahu-Hawaii circuits have proved a very valuable asset to the interisland radio network. With all sites supervised from Honolulu, and with complete control of all equipment accomplished from Honolulu, major savings in operating and maintenance costs have been achieved.

References

1. The Interisland Radio Network of the Hawaiian Telephone Company, W. B. Farinon. *AIEE Transactions*, vol. 74, pt. I, Sept. 1955, pp. 529–36.

2. Engineering Problems and Performance Record of the Hawaiian Interisland Radio Network, W. B. Farinon, A. C. Walker. *Ibid.*, Nov. 1955, pp. 571–76.

The Ticketfax System

A. S. HILL
ASSOCIATE MEMBER AIEE

Synopsis: In an improved system for the control of reserved Pullman accommodations, the car diagram has been replaced by a set of coupons, each representing a unit of space in the car. Breaking up the car diagram into its component parts represented by separate coupons eliminated a bottleneck impeding the operations of determining space availability and effecting its sale. Data on a space coupon are rapidly reproduced by facsimile and without handwriting onto a preprinted Pullman ticket form, either locally or over some distance, thereby greatly reducing the time required to effect the sale of space to a waiting customer. Ticketfax transmitters and recorders were specially designed by Western Union to meet the unique requirements of this phase of the improved system.

Current General Practices

WITH the car diagram system the procedure of handling reservations and sales of space is described briefly as follows:

In making a reservation the customer or a ticket office clerk telephones the reservation bureau which assigns the space and records the name of the customer on the car diagram and specifies the limit of time in which it must be purchased. Within the specified time limit the customer or his messenger goes to the ticket office to pick up the ticket.

In making the actual sale, the ticket seller calls the reservation bureau confirming the space assignment. He fills in a ticket, which is prenumbered and color coded for the type of accommodation, with the designations of the space assigned and the car in which it is located, the price, tax, and total cost, data and time of train departure, and in some cases the ON and OFF stations. The reservation bureau records in the car diagram the number of ticket under which the accommodation is sold, the ticket seller's identification number, and ON and OFF stations, and the make-down information, i.e., the number of persons for whom the space is to be prepared. With the car diagram method of space control it is difficult to obtain complete information on space availability especially during peak sales activities.

Improvements Sought

The Pennsylvania and New York Central systems desired an improved method of reserved space control and sale. All phases of travel information, space reservation, and ticket selling were examined with the view to devising means whereby reserved space would be brought under complete control and all operations involved would be more closely co-ordinated and simplified. The two main objectives were to make it easier for the customer to secure reservations and to obtain tickets, and to make possible more nearly complete sale of accommodations.

A study of practices disclosed that the car diagram, which tied up all units of space in a particular car in one record, was the bottleneck which impeded all operations involved in reserving and selling space and making space available for resale released through cancellations. The scheme by which the car diagram method is improved upon and the sale of reserved space simplified and speeded up is basically quite simple. Instead of the car diagrams, there is a coupon to represent each unit of reserved space available on all trains as far into the future as

---

Paper 55-501, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 28, 1955; made available for printing May 6, 1955.

A. S. Hill is with the Western Union Telegraph Company, New York, N. Y.

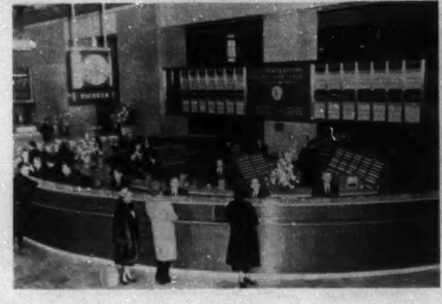


Fig. 1 (left). New ticket office in the Pennsylvania Railroad 30th Street, Philadelphia, Station, showing coupon and files availability board in the background

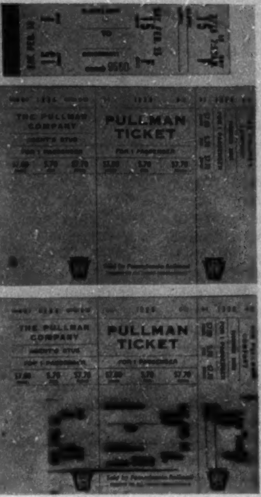


Fig. 2 (right). Top—Space coupon Middle — Pullman ticket form Bottom — Pullman revenue ticket

necessary. A coupon is coded with all pertinent information about the space it represents such as location in the car, connecting bedrooms, crosswise and lengthwise berths, window or aisle seat, etc. Thus replacing the diagram by coupons representing its component parts permits separation of clerical recording functions, ready accessibility to available space, and fast processing of cancellations. Most importantly, the data on a space coupon can be rapidly reproduced by facsimile onto a ticket at an outside agency where a customer is waiting. Furthermore it permits reproductions of data on reserved space coupons by facsimile onto tickets in offices of customers.

Early in 1953 the coupon method for control and sale of space was put into operation by the Pennsylvania Railroad at Pittsburgh, Pa. Many of the equipments employed in that installation, including those supplied by Western Union, were hurriedly devised modifications of machines originally designed to serve entirely different purposes. Nevertheless, the results of the operation of the installation were so unmistakably beneficial that the railroad group planned a more extensive application of the system.

During January 1955 system installations were made for the New York Central Railroad at their Cleveland terminal station and for the Pennsylvania Railroad at their 30th Street Station, Philadelphia, Pa.; see Fig. 1. In these installations two types of Western Union Intrafax equipment provide the special communication features vital to the operation of the system.

Specially designed Ticketfax machines, which are fully described in companion papers[1,2] provide the high-speed transmission necessary between the reservation bureau and the ticket clerks. (The name "Ticketfax" has been given to facsimile transmitters and recorders designed by Western Union to meet the unique communication requirements of an improved method of controlling and selling reserved accommodations on railroad passenger trains.) In addition, standard speed Western Union Intrafax message-type equipments, slightly modified, are employed on the direct circuits to customers' business offices.

The Improved System

Starting with the space coupon, the important features of the system are described here, with emphasis on the communication requirements served by the facsimile equipments.

The space coupon is of white cardboard $4^1/_2$ by $1^1/_2$ inches. While the coupon, Fig. 2, always remains in one piece, it is printed in three sections containing respectively the information required by the agent selling space, the Pullman conductor on the train, and the customer before and while on the train, and later for expense records.

Across the front and back of one end of the coupon, there is a 3/8-inch wide strip of the color indicating the type of accommodation it represents. Printing on the coupon is in black. The designations of the car and space unit are printed in large bold type. The stations between which the car travels and other data are printed in smaller type. There are spaces for filling in ON and OFF stations when the space is sold for trips starting and/or ending at intermediate points. The control center, a subdivision of the Ticket Sales and Service Bureau, has custody of coupons.

The coupons are completely printed in advance in large quantities, except for dates and departure times, and the printer assembles them in car sets. Before being placed on sale, coupons are imprinted with the use date in each of the three sections plus the train departure time in the customer's section. The tiered coupon files for all space on all trains weeks in advance are convenient to all control center personnel. By a file-dating arrangement and the coupon color-coding, the desired coupon can be readily located and extracted.

The coupon, which is the sole control of the space it represents, does not leave the custody of the control center. Reposing openly in the for-sale file it represents a space available for immediate sale or reservation. When reserved it is enclosed in a color-coded paper jacket but remains in the file. When sold it is relayed to a sold file from which it can be quickly recovered and restored to the for-sale file in case of cancellation.

THE BLANK PULLMAN TICKET FORM

Ticket sellers at the main ticket office, satellite stations, city ticket offices and agencies served by the control center are supplied with blank Pullman ticket forms printed on Teledeltos* facsimile recording paper; see Fig. 2. The forms are printed with colored inks for accommodations type coding. A ticket form is in three sections separated by perforations. The printed heading of each section includes the price, tax, and total charge for the space which it will represent when sold. Although the ticket forms are serially numbered and must be accounted for, they are not negotiable until processed as described further on. Blank Pullman ticket forms supplied to private business subscribers are very similar to those described in the foregoing except that the printing does not include the price.

---

* Registered trade-mark of the Western Union Telegraph Company.

THE PULLMAN SPACE REVENUE TICKET

When the information in each of the three sections of a coupon is reproduced respectively under the printed headings of the three sections of a properly priced and color-code printed Pullman ticket form, a valid Pullman reserved-space revenue ticket results; see Fig. 2. The reproduction is accomplished by the Ticketfax machines and by telegraph message-type Intrafax machines in the case of private customer service.

RAPID FACSIMILE TRANSMISSION

As stated before, one of the main objectives of the system is to make it easier for the customer to procure his Pullman space ticket. This includes reducing his waiting time to the lowest possible minimum. Since, as in the older system, the Pullman ticket is not produced until the customer calls for it, it then must be produced as rapidly as possible. Reproduction by the facsimile method is accomplished by optically scanning, successively, in a predetermined pattern, minute areas of the subject copy, and reproducing the optical effects of these minute areas in the same pattern. The time required to transmit a document is related to its size. In this system the size of the coupon was held to the absolute minimum possible for the purposes to be served. Ticketfax machines operating at 30 scanning strokes per second and reproducing the coupon data in about 6 seconds are employed for transmission from the control center to ticket clerks across the counter and at satellite stations several miles away. Machines operating at six scanning strokes per second, requiring about 30 seconds to reproduce the coupon, are employed on circuits connecting more distant satellite stations.

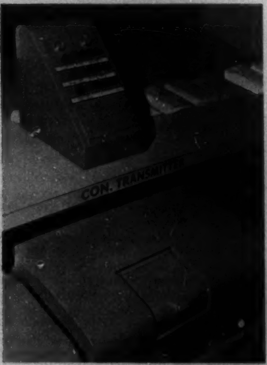


Fig. 3. Line select push-button unit

In addition to coupons, other forms are transmitted over the Ticketfax circuits. Satellite stations transmit order and cancellation forms to the control center. The control center sends space availability forms to satellite stations. Since at satellite stations coupons are received on printed Pullman ticket forms and other messages are received on unprinted recording paper, two different calling signals are provided so that the called station will know what recording form to use in answering a call.

Operation

It is deemed sufficient for the purposes of this paper to discuss in detail only that phase of the system by which reserved Pullman space tickets are provided to satellite stations in the Pennsylvania Railroad installation. This phase includes all of the operations involved in providing tickets to ticket sellers in the main ticket office adjacent to the control center.

The ticket sale transaction starts at the satellite station. A customer wants a certain type of accommodation on a particular train between two particular stations for one or more persons. The station ticket seller fills in an order form with the necessary information including the name of the customer and the ticket seller's individual code symbol. He places the form on the transmitter and presses the message-send start button.

As the satellite station ticket seller presses the start button, a relay is actuated in the control center concentrator causing the lowest numbered idle and readied recorder to be connected to the line. The recorder starts, phases with the calling station transmitter, both machines start scanning, and transmission is completed in 6 or 8 seconds. End-of-message signals are actuated on the machines at both ends of the circuit.

At the control center a space distributor attends the recorder showing an end-of-message signal. He removes the recorded message from the recorder and, if he is able to identify the sending station from the message, he extinguishes the end-of-message signal and the recorder is disconnected from the circuit. At the satellite station the ticket seller attends his transmitter by extinguishing the end-of-message signal and removing the transmitted request form from it; the line is automatically transferred from his transmitter to his recorder. The ticket seller then selects from a rack a Pullman ticket form printed in the color representing the type of accommodation requested and of the proper value for the number of persons to occupy the space, in anticipation of a call from the control center to send him the space.

At the control center the space distributor takes from the file a coupon representing a unit of the type of accommodation requested on the desired train. He marks it with his code symbol and places it on an idle transmitter, presses the coupon-send start button and the proper line-select button, Fig. 3, to cause the concentrator to connect his transmitter to the satellite station line.

At the satellite station the calling buzzer sounds and the coupon-receive lamp glows indicating to the ticket seller that he is to receive the space requested. He places the Pullman ticket form on his recorder and presses the start button. His recorder phases with the control center transmitter and transmission proceeds. At the end of transmission the end-of-message signals are again actuated at both ends of the circuit.

At the satellite station the ticket seller

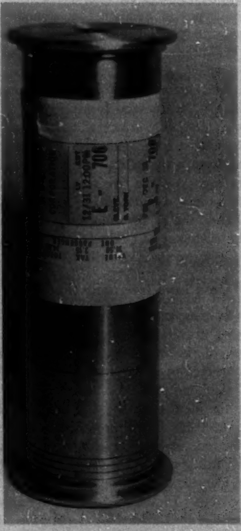


Fig. 4. Special transmitter drum for sending coupons to industrial customers

extinguishes the end-of-message signal, removes the completed Pullman space revenue ticket form, detaches the agent's stub section, and sells the remaining two attached sections to the waiting customer. At the control center the clerk extinguishes the end-of-message signals on the transmitter, causing it to be disconnected from the line. He inserts the transmitted coupon into an automatic stamp which puts on the back the word "sold" and the time and date of sale. The coupon is then relayed to the sold file. Ordinarily the time required for the complete transaction described is measured in seconds.

DESIRED SPACE NOT AVAILABLE

If in the foregoing transaction the type of accommodation requested had not been available, the space distributor would have filled in an availability form to show the types of accommodation available on the desired train and transmitted it to the requesting station in lieu of a coupon. In this case he would have started his transmitter by pressing the message-send start button. At the satellite station the calling signals would have included the glowing message-receive lamp. The ticket seller at the satellite station would have placed a blank sheet of recording paper on his recorder instead of the printed ticket form. Assuming that the waiting customer would have indicated which of the types of available accommodations he would accept, the transaction would have been completed on this basis as before.

The customer purchasing accommodations in the transaction just described presents the two attached sections of his Pullman space revenue ticket to the Pullman conductor on his train. The conductor separates the two sections, retaining the proper section for his purposes, and returns the smaller remaining section to the customer as evidence of his right to occupy the space described thereon.

STATION LINE IDENTIFICATION

As mentioned in describing the foregoing transaction, an incoming call on any of the many lines is automatically connected to an idle and readied recorder. There is no indication on the recorder as to the identity of the line to which it is connected. If the prescribed routine is followed, and equipment units involved perform properly, the sending station can be identified by the recorded copy of his message. In the event that a received message does not include the identification of the sending station, the line-identification push button is held depressed to light the lamp in a line-identification panel which is associated with the line to which the recorder is connected; the line-identification panel is situated to be visible from all concentrator recorder positions.

SERVICE TO PRIVATE BUSINESS CUSTOMERS

Under the coupon system, as mentioned in the foregoing, Pullman tickets can be delivered by wire to the offices of private business customers. Arrangements can be made to furnish such customers with supplies of rail fare tickets on a deferred payment plan. Western Union telegraph-type Intrafax equipments with minor modifications are employed in this phase of the system. Desk-Fax transceivers are employed by the customers for ordering Pullman accommodations and receiving Pullman revenue tickets.

At the Ticket Sales and Service Bureau there is a concentrator comprising a turret with a capacity of 30 lines and consoles accommodating four removable drum transmitters and four continuous paper-roll page recorders.[3-5]

In sending to customers two types of transmitter drums are used; one to send messages and forms other than coupons, the other to send coupons with price figures. Around the latter type of drum, steel bands form guides into which the space coupons and a price strip are inserted. One of these latter type of drums, Fig. 4, is prepared for each subscribing customer. The coupon and price strip do not completely encircle the drum. On the otherwise unused scanned area of each drum is permanently imprinted the name of the customer for whom the drum is exclusively used. The customer's name is transmitted with each coupon and is recorded on two sections on his ticket along with space and price data. At the same time that the ticket is being delivered in the customer's office, one of two monitor recorders in the control center produces, for billing purposes, a facsimile reproduction of the transmission including the customer's name.

Availability Board

Many very important features of the over-all ticketing system do not fall within the scope of this paper and have been omitted. An exception is made in favor of one such feature, namely, the manually operated, mechanical availability board, becuase of its helpfulness to both railroad people and customers.

The availability board is a framework divided into 36 rectangular sections each in the order of 18 inches wide by 2 feet high. Each section shows the up-to-the-minute availability of the type of reserved accommodations on one regular train which is prominently designated at the top of the section by name, number, and departure time. The area in the section under the heading is divided into 49 squares of about 2 by 2 inches, arranged in seven vertical columns for 7 days starting with the current day and seven rows across for types of accommodations carried up to a maximum of seven.

The small squares show a type of accommodation on a white background if there are many units of that type available. When there are only a few units left of the type of accommodation represented, the background of the square is changed to yellow. When the type of accommodation represented is completely sold, the square becomes blank. With the aid of a long-handled tool, changes in a small square are made by manipulating a small lever associated with it. The vertical columns of squares can be removed as units for daily rearrangement.

The entire board is suspended above, and spans across the entire office, Fig. 1, so as to be visible to the ticket sellers and customers from the front, and to telephone clerks and other personnel from the rear. A glance at the board gives a complete picture of accommodations availability on all trains for the next 7 days.

Conclusion

The systems at Cleveland and Philadelphia have been in operation for a few months. Performances are being closely observed and work is in progress on minor changes which experience indicates to be desirable. The Ticketfax machines have performed quite satisfactorily. The main objectives of the system have been realized to a great degree and it is expected that installations in other large city terminals will follow.

References

1. TICKETFAX ELECTRONICS, C. Jelinek, Jr. *AIEE Transactions*, vol. 74, pt. I, Nov. 1955, pp. 565–68.

2. TICKETFAX TRANSMITTERS AND RECORDERS, D. M. Zabriskie. *Ibid.*, Sept. 1955, pp. 536–41.

3. A VERTICAL DRUM TELEFAX TRANSMITTER, J. H. Hackenberg. *Western Union Technical Review*, New York, N. Y., vol. 4, no. 1, Jan. 1950.

4. THREE-STYLUS FACSIMILE RECORDER FOR CONCENTRATOR SERVICE, F. G. Hallden. *Ibid.*

5. THIRTY-LINE CONCENTRATOR FOR PRIVATE FACSIMILE SYSTEMS, H. H. Haglund. *Ibid.*, vol. 6, no. 4, Oct. 1952.

Phase Locking of Switching-Transistor Converters for Polyphase Power Supplies

A. G. MILNES
MEMBER AIEE

SWITCHING transistors may be used in conjunction with saturable-core output transformers to convert direct-voltage inputs into alternating voltages of nearly rectangular wave form.[1] The output impedance is low, the over-all conversion efficiency is high, and (even in the present state of power-transistor development) output powers of tens of watts are available. Such transistor converters are therefore being considered as supply sources for magnetic amplifiers, a-c servomotors and other components used in low-power servocontrol systems. For example, in air-borne military applications, transistor power supplies may replace small rotating machine inverters, which at present produce 400-cycle-per-second supplies from the aircraft battery system, with advantages in reliability and weight.

From a normal single-phase sinusoidal supply a 2-phase servomotor may be operated by supplying one of the phase windings through a capacitor to give the required field rotation. However, for a rectangular output voltage wave form this method is not satisfactory and, therefore, consideration is given to phase locking of two transistor converters in a prescribed 90-degree (°) phase sequence. The action of the proposed circuit is studied under both steady-state and transient conditions. Since the arrangement may be designed to provide any desired angle of phase lock, three converters may be set in 120° phase sequence to give a 3-phase system if required. These techniques for linking converters to form polyphase systems are likely to enhance considerably their attractiveness for power supply purposes.

Basic Converter Circuit

The basic converter circuit is shown in Fig. 1(A) and some relevant transistor characteristics are given in Fig. 1(B). Since the circuit action has been discussed by Royer,[1] only a brief account need be given here for the purposes of orientation.

Consider the half-cycle when transistor A is conducting and the battery voltage E is applied to winding 2. This voltage causes flux increase in the transformer core and induces a voltage across winding 1 which provides base current for the transistor. The characteristics in Fig. 1(B) are for a PNP power transistor which is seen to require a base current of about 55 milliamperes for a collector current of 600 milliamperes under room-temperature conditions. The required emitter-base voltage is then of the order of 1.2 volts, so winding 1 normally has fewer turns than winding 2. During the period of flux increase, the induced voltage in winding 4 is in the sense that maintains transistor B nonconducting since the base is positive with respect to the emitter. Because of the induced voltage in winding 3, the total emitter-collector voltage appearing across transistor B in the cutoff condition is equal to twice the battery voltage.

At the end of the half-cycle, the flux in the transformer core reaches the positive saturation level $+\Phi_s$ and the drive and holdoff voltages in the base circuits of transistors A and B respectively then collapse. This permits the two transistors to interchange action so that the input voltage is applied to winding 3 and the core flux then proceeds towards negative saturation. If the windings 2 and 3 have N turns each, the voltage-flux relationship is

$$\int_0^{1/2f} E\,dt = \int_{-\Phi_s}^{+\Phi_s} N\,d\phi$$

whence the frequency of oscillation is given by

$$f = \frac{E}{4N\Phi_s}$$

Thus, for a given saturable transformer the frequency of the alternating output is proportional to the input voltage. If the input voltage is increased, the output amplitude increases proportionately, but the period of oscillation decreases (i.e, the frequency increases) and the voltage-time integral of the output is unchanged.

In some aircraft applications the battery system available may vary as much as from 21 to 29 volts. For magnetic amplifier supply purposes a gate voltage of constant voltage-time integral is advantageous since it corresponds to constant excitation of the cores in terms of flux swing. The nearly rectangular wave form of the converter output is not objectionable for such supply purposes.

In the experimental work to be described, power transistors were used but no attempt was made to operate them at or near their limits of performance since cores of suitably low winding resistances were not available and attention was concentrated more on the principle of phase locking. The maximum quoted rating for the transistor used are a collector voltage of 60 volts and a collector current of 800 milliamperes. Thus, for a nominal 24-volt battery system input, the output power can be of the order of 20 watts. A collector dissipation of 5 watts is the maximum recommended with the mounting base maintained at 70° centigrade; in a test on a unit mounted in the center of a 6-inch by 6-inch metal plate, a thermocouple under the mounting stud of the transistor indicated a temperature rise of 40° centigrade in still air for 5-watt collector dissipation. This dissipation limit is represented by the curved boundary line in Fig. 1(B). The straight line PQ represents a possible load line and this can pass above the dissipation boundary without danger of overheating of the transistor provided the switching action between points P and Q occurs rapidly.[2] This circuit switching time is likely to be influenced by saturation inductances of the circuit as well as by fundamental transistor properties, but designs which are capable of operating at frequencies up to 12 kc per second or more have been constructed.[1,3]

In the basic converter circuit a high-frequency oscillation, corresponding to partial flux swing in the transformer, may occur if the load impedance is decreased to require a peak collector current in excess of that permitted by the base current drive. If, further, the output is inadvertently short-circuited, both transistors

---

Paper 55-543, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 28, 1955; made available for printing May 3, 1955.

A. G. MILNES is with the Carnegie Institute of Technology, Pittsburgh, Pa.

The author wishes to acknowledge his indebtedness to Prof. L. A. Finzi and colleagues at Carnegie Institute of Technology for stimulating this work and for the facilities provided under the United States Research Fellowship held.

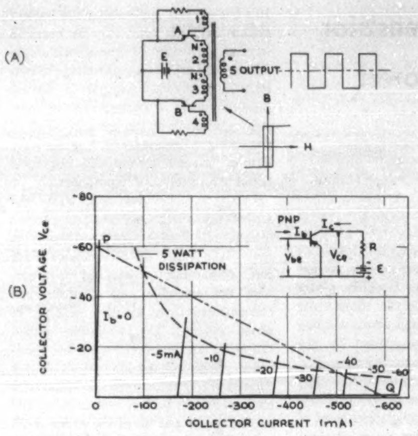


Fig. 1 (left). Switching transistor d-c to a-c converter

A—Basic circuit with PNP transistors
B—Collector characteristics of power transistor

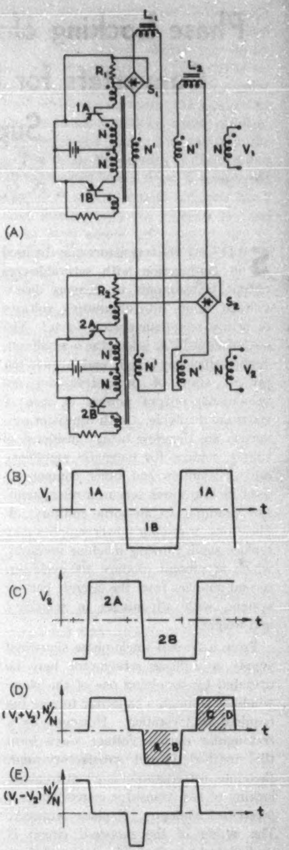


Fig. 2 (right). Quadrature phase locking of transistor converters

A—Circuit arrangement
B through E—Voltage wave forms, without phase-locking

become nonconducting and no lasting damage occurs. This is an advantage of the basic circuit that is fortunately retained in the phase-locking circuit to be discussed in the following.

Quadrature Phase-Locking Circuit

The phase difference of two similar wave forms may be detected by considering the sum and difference of the two output voltages, and this principle is used in the circuit shown in Fig. 2(A). The sum of the two voltages, proportional to the converter output voltages, is applied to the circuit comprising the saturable inductance L_1 and a full-wave rectifier bridge which injects a voltage across the resistor R_1 in the base circuit of transistor 1*A*.

This voltage sum has the wave form in Fig. 2(D) if a phase lag of less than 90° exists between V_1 and V_2, Fig. 2(B) and (C). The saturable reactor L_1 is proportioned so that its full flux swing from positive to negative saturation levels corresponds to the shaded voltage-time areas *A* or *C* in Fig. 2(D). Hence the surplus voltage-time areas *B* or *D* appear across the rectifier bridge and so across the resistor R_1 in a sense tending to make the base positive with respect to the emitter and, therefore, to cut off the transistor conduction. The voltage pulse *B* is not of consequence since the transistor 1*A* is nonconducting when it occurs, but pulse *D* begins when the transistor is conducting and hence shortens the conduction period. If the saturable reactor L_1 has a voltage-time-area rating corresponding to the voltage sum when the phase difference is 90°, the wave form of converter 1 will therefore be advanced relative to that of converter 2, thus bringing the waves more into quadrature.

The circuit containing the difference of the two converter voltage becomes operative if the phase lag exceeds 90° since then reactor L_2 saturates and advances the wave form of converter 2 to provide the required correction. Thus, if reactors L_1 and L_2 are equal and proportioned to have a voltage-time-area rating corresponding to half of that of the converter saturable transformers (based on N' turns), the circuit shown in Fig. 2(A) results in quadrature phase locking with V_2 lagging 90° on V_1. Variation of input battery voltage does not disturb the phase-locking action.

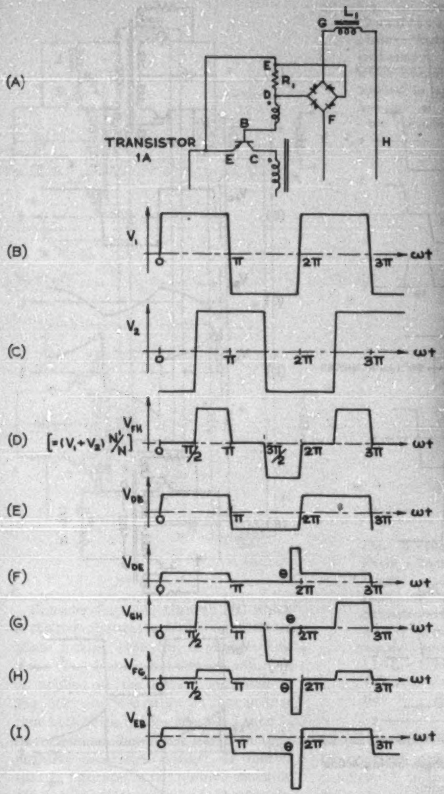
The account of the circuit action given in the foregoing is basically correct but a shortened conduction period in one-half cycle represents a reduced flux swing in the main transformer core and this involves a short conduction period in the next half-cycle. Part of the phase-correction action is lost on this account but a step-by-step process follows until phase locking is achieved after several cycles.

The circuit wave forms when two converters are locked in quadrature are of interest and are given in Fig. 3 for the simplifying assumptions that:

1. The impedance of reactor L_1 when not saturated is large compared with R_1, but is negligibly small when saturated.
2. The winding resistances are small so that the current-resistance voltage drops across them may be neglected.
3. The rectifier bridge is ideal, i.e., has negligible forward resistance and negligible leakage current.

Fig. 3(A) shows the portion of the circuit, of Fig. 2(A), that is under consideration and Figs. 3(B) and (C) represent the two converter wave forms of output voltage with 90° phase shift. The voltage sum applied to the locking circuit is therefore as shown in Fig. 3(D). The voltage across the drive winding in the base circuit of the transistor is represented by Fig. 3(E). During the conducting period it is assumed that about half the drive voltage appears across the emitter-base terminals and that the remainder appears across the resistor R_1 as voltage drop produced by the flow of base current; see wave form in Fig. 3(F) between 0 and π. This voltage drop reduces the amplitude of the voltage applied across

Vol. 7
Pt. 1
1955



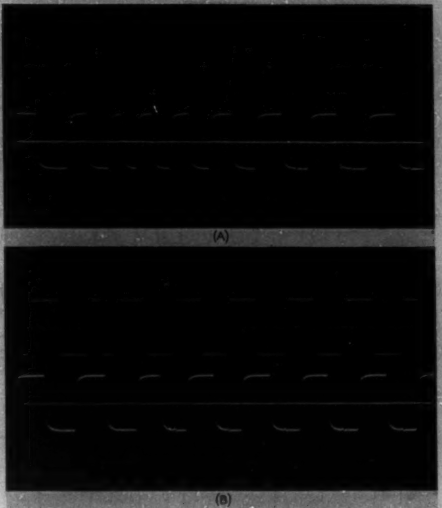


Fig. 3 (left). Voltage wave forms for the quadrature phase-locked condition

Fig. 4 (above). Phase-locking transient oscillograms

A—90° lag achieved from the in-phase condition
B—90° lag achieved from the 90° lead condition

the saturable reactor L_1 and therefore V_{GH} in wave form 3(G) is less than V_{FH} during the period $\pi/2$ to π. It is assumed that the voltage-time area of V_{GH} between $\pi/2$ and π corresponds to full flux swing in the saturable reactor L_1. In the next half-cycle no base-emitter current flows and so the full locking voltage V_{FH} appears across the saturable reactor from the position $3\pi/2$ to a saturation instant θ (earlier than 2π because of the greater amplitude of the voltage applied to the reactor). Between θ and 2π, since the reactor is saturated, the voltage V_{FH} appears across the rectifier bridge and so across the resistor R_1 as a voltage pulse. This pulse occurs merely as additional switch-off voltage, Fig. 3(I). during a period when the transistor is (in any case) nonconducting. It is, however, a prominent feature of the circuit voltage wave forms and liable to misinterpretation as a phase-locking or transistor switching pulse.

One feature of importance brought out by the foregoing discussion is that the required rating of the saturable reactor is not exactly half the voltage-time rating of the converter (based on N' turns) as proposed earlier, but somewhat less because of the voltage produced across R_1 by the flow of the base current. The voltage drop across R_1 results in the pulse widths of the saturable reactor voltage being different in the two half-cycles, and under transient conditions (discussed in the next section) this effect contributes to the phase-locking action.

Experiments with the phase-locking circuit arrangement of Fig. 2(A) showed excellent phase holding if the load on either phase was varied or if the input voltage was changed in the range of 20 to 30 volts. If resistance was inserted in the primary supply lead of one converter to cause unbalance that would normally give a difference in switching rates of the two units, the phase locking still maintained control. Although two separate battery supplies are shown in Fig. 2(A), this is only for convenience in the layout of the diagram and in the experiments the two converters were operated from the same battery system. The converters used H_2 (Minneapolis-Honeywell) transistors and the frequency was about 90 cps for a 24-volt input.

Transient Behavior When Phase Locking Is Applied

To examine the process of phase locking, numerous observations were made of the voltage output wave forms of two converters brought into phase sequence by sudden switching in of the phase-locking circuits at the points S_1 and S_2 in Fig. 2(A). Correct sequence locking was obtained for all phase positions of the two wave forms at the instant of switching. The two converters were well balanced

ol. 74
. 1
55

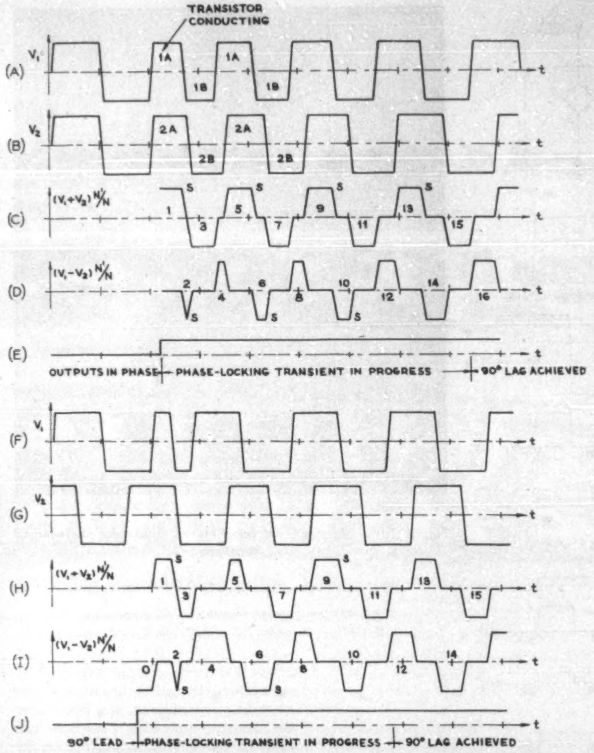


Fig. 5. Theoretical transient wave forms on application of the quadrature phase-locking action

A through E—90° lag of V_2 achieved from initial in-phase condition
F through J—90° lag of V_2 achieved from initial 90° lead condition

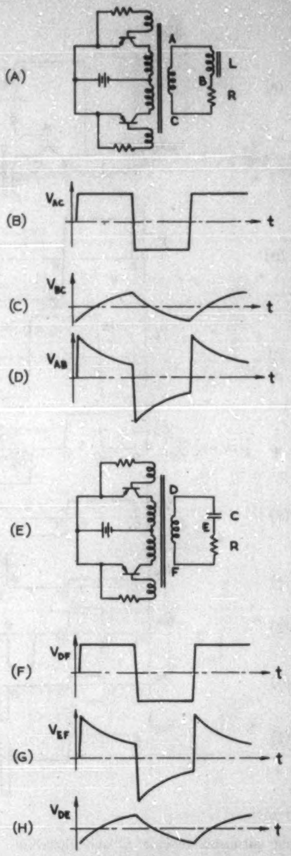


Fig. 6. Converter behavior with inductive or capacitive loads

A through D—Inductive load
E through H—Capacitive load

before the switching operation to ensure that slipping because of frequency difference was not an important factor in the locking action observed. Oscillograms are given in Fig. 4(A) and (B) to show locking occurring from an initial in-phase condition and from a 90° lead condition. The main two traces represent the converter output voltages V_1 and V_2, and the third trace in each case is a marker to indicate the instant of switching in the phase-locking circuits. From the in-phase initial condition, the prescribed 90° lag sequence is achieved in Fig. 4(A) after about 8 cycles, and from the 90° lead condition of Fig. 4(B) the action is even quicker. Transients occurring when the input battery voltage was applied to the two converters complete with phase-locking circuits were also observed and found to be comparable.

If attempts are made to explain the phase-locking action on a fully idealized basis, the final steady-state behavior theoretically deduced depends very much on a combination of factors such as:

1. The phase-sequence condition existing prior to switching.
2. The actual instant in this sequence at which switching is performed.
3. The flux conditions of the phase-locking reactors L_1 and L_2 at the switching instant.

For certain combinations of these factors, phase locking in quadrature can not be expected from the circuit for fully rectangular wave forms of output voltage, and it is necessary to make a more realistic approach by recognizing that the switching process of a converter takes a discernable time to develop, although typically not more than, say, 1/20th of the half-cycle period. The saturation inductances of the converter transformer windings may determine this period of switching since in purely resistive circuits the switching times of transistors are low, say, of the order of tens of microseconds. In Fig. 5 some theoretical transient wave forms for phase locking with the circuit of Fig. 2(A) are given on the assumption that the converter wave forms are slightly trapezoidal.

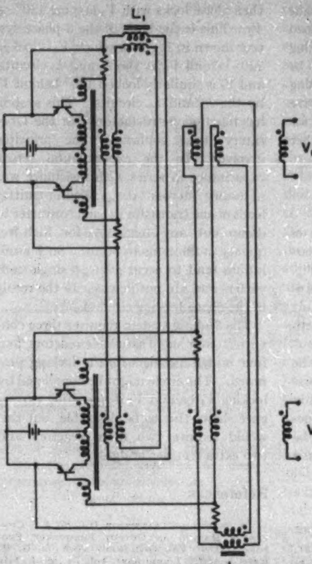


Fig. 7 (left). Quadrature phase-locking arrangement in which control is applied to each transistor of the two converters

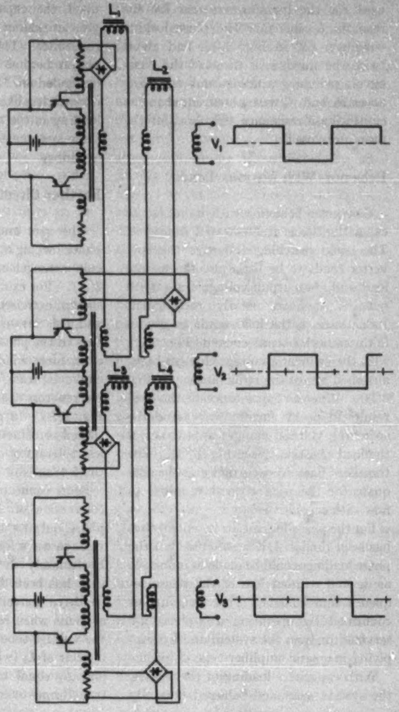


Fig. 8 (right). Phase-locking circuit arrangement for three converters to provide a 3-phase supply system; $L_1=L_3=1/2L_2=1/2L_4$ in saturation rating

Consider Fig. 5(A) through (E) which is the hypothetical action for quadrature phase locking from the in-phase condition. The locking circuits are regarded as applied at the instant indicated in Fig. 5(E) and it is supposed that at that time both of the locking reactors L_1 and L_2 have residual fluxes just less than their negative saturation values. It follows that L_1 saturates in the positive direction after about 1/4th cycle and thus transistor 1*A* is switched off after only about half its normal conducting time. This produces voltage across L_2 and this saturates in the negative direction almost immediately and switches off transistor 2*A*. Since 1*A* has conducted for only about half its normal period, the main transformer flux swing has been curtailed and so the next switching action is by 1*B* after an equal period.

The succeeding action may be followed step by step in the same way. The sum and difference wave forms for the locking circuits are shown in Fig. 5(C) and (D), where the letters *S* indicate the times when the circuits cause transistor switching. In Fig. 5(D), it will be noticed that pulse 6 is greater in width than pulse 4, and similarly for pulses 10 and 8, or 14 and 12. This occurs because account is being taken of the voltage drop across the resistor R_2 in the base circuit of transistor 2*A* (as discussed with reference to R_1 in the previous section). This contributes significantly to the phase-locking action, which is a gradual process comparable with that shown in the oscillogram of Fig. 4(A).

Hypothetical wave forms for phase locking from the 90° lead condition are given in Fig. 5(F) through (J). Here it is assumed that initially the flux in L_1 is near the negative saturation value and that for L_2 the flux is near positive saturation. It will be seen that the theoretical switching action that results closely parallels that shown in the oscillogram of Fig. 4(B). For the 90° lead condition, the theoretical possibility exists that the initial flux conditions in both L_1 and L_2 may happen to correlate with the instant of switching in such a way that equilibrium exists. However, stability with 90° lead of V_2 was never observed in practice although attempts were made to achieve it in locking tests on converters carefully balanced to be free from phase slip. In normal practice, converters are not likely to be perfectly balanced and any phase-slip tendency makes stability in the 90° lead condition, instead of the prescribed 90° lag sequence, virtually impossible.

If the reactors L_1 and L_2 are made with a voltage-time rating smaller than appropriate to the converter transformer rating, the system frequency is controlled by the reactors and is higher than it otherwise would be because the main transformers do not experience complete flux swings. In single converters, control of flux swing in the transformer by bias techniques as a means of increasing frequency has been discussed by Van Allen.[4] On the other hand if the reactors L_1 and L_2 are too great in voltage-time rating, the locking does not accurately maintain the quadrature condition. With properly proportioned reactors, the locking is held to within ±5° of the 90° lag position for all normal variations of phase loading and input voltage.

In the converters used for the experimental work, cores of Hipernik V were used for the transformers and for the reactors L_1 and L_2. The phase-locking windings, (N' in Fig. 2(A), had about twice the number of turns of the base-circuit switching windings and the resistances R_1 and R_2 were about equal to the emitter-base resistance (25 ohms for the transistors used).

Behavior With Various Loads

Converter behavior with inductive or capacitive loads is illustrated in Fig. 6. The basic switching action of the converter tends to be little affected by the load and the output voltage wave form remains therefore nearly rectangular. Inductance in the load tends to give a fairly smooth output current, Fig. 6(C), with the switching voltage changes being absorbed across the inductance; see Fig. 6(D). Whereas capacitance in the load results in peaky current waves and the switching voltage changes appear across the load resistance; see Fig. 6(G). The transistor base-drive current must be adequate for the peak current required to flow with capacitive load.

For the phase-locking arrangement that has been proposed, it was found that the phase loadings could be made as inductive as desired without loss of the phase sequence and control. This was further confirmed by operating a 2-phase a-c servomotor from the system and by supplying magnetic amplifier type of loads.

With capacitive loading of the phases, the system examined behaved normally until the capacitive reactance became greater than or comparable with the load resistance. Thereafter, the phase-locking action became eratic and could not be depended on. Since this type of loading is much less likely to occur than inductive loading in the applications envisaged for these systems, the matter was not further examined.

Further Circuit Developments

The sum and difference technique of phase locking may be applied in other circuit arrangements besides that of Fig. 2(A). For example, in Fig. 7, control is shown exercised on all four transistors and this arrangement requires no rectifiers in the phase-locking circuits. However, more windings are required on the converter transformers and on the saturable reactors than for the arrangement in Fig. 2(A). In practice this circuit functioned satisfactorily but seemed to be less tolerant of converter unbalance than the 2-transistor controlled circuit.

Scott-connected transformers for the conversion of 2-phase supplies into 3-phase outputs are not suitable for input voltage wave forms that are rectangular. Fortunately the phase-locking principle that has been discussed may be applied to three converters to produce 3-phase systems when required. If in Fig. 2(A) the voltage-time rating of L_1 is made half of that of L_2 (with the total rating of L_1 and L_2 equal to that of the converter transformer over a half-cycle), the circuit then phase locks with V_2 lagging 120° on V_1. This is the basis of the 3-phase system shown in Fig. 8 in which V_2 is locked 120° behind V_1 by the L_1 and L_2 circuits, and V_3 is similarly locked 120° behind V_2 by the L_3 and L_4 circuits. This system has functioned satisfactorily in the laboratory, thus confirming the principle involved. In the experimental setup, capacitance in series with resistance was necessary across the collector-emitter leads of one transistor in each converter to damp out any tendency for high-frequency oscillations to occur. Such oscillations tend to occur even in single converters and are not necessarily the result of the phase-locking circuits.

The 3-phase system requires three converters, four small saturable reactors, and four bridge rectifiers for the locking proposed. The circle might be completed by locking V_1 behind V_3 if operating experience shows this to be desirable, but this would require two more reactors and two extra rectifier bridges.

References

1. A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage, G. H. Royer. *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 322–26.

2. Junction Transistors Used as Switches, R. L. Bright. *Ibid.*, Mar. 1955, pp. 111–21.

3. A New Self-Excited Square-Wave Transistor Power Oscillator, G. C. Uchrin, W. O. Taylor. *Proceedings*, Institute of Radio Engineers, New York, N. Y., Jan. 1955, p. 99.

4. A Variable Frequency Magnetic-Coupled Multivibrator, R. L. Van Allen. *AIEE Transactions*, vol. 74, pt I, July 1955, pp. 356–61.

Magnetic Amplifiers in Bistable Operation

L. A. FINZI
MEMBER AIEE

G. C. FETH
STUDENT MEMBER AIEE

MAGNETIC amplifiers in bistable operation behave substantially as a latch-type relay.[1] Two widely different output levels can be obtained stably, one close to the quiescent and the other close to the saturated output. Over a certain range of the applied signal the one or the other condition is found to occur, depending on whether this range has been entered by increasing or decreasing values of signal.

The amplifiers considered in this paper are those in which bistable operation (called also "flip-flop" or "snap" operation) results from overregenerative magnetomotive forces (mmf) introduced intentionally to modify the transfer characteristic of conventional self-saturating amplifiers. Aspects of the mechanism of transition from the one to the other output level are considered, and criteria are offered for predictions of the time of response.

Mmf Forces of Overregeneration

The familiar approach is chosen in which mmf are viewed as the primary factor governing the angle of firing of the amplifier under conditions of comparatively low suppression index.[2] With this choice the practical influence of the magnetic properties of the cores of the bistable amplifier can be accounted for in realistic terms readily available to the designer, i.e., in terms of transfer characteristics, predicted or experimentally known, of stable amplifiers without overregeneration.

The cores of a magnetic amplifier operating at a given angle of firing α, and thus at a certain output voltage, describe a minor hysteresis loop in a way which

---

Paper 55-539, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 28, 1955; made available for printing May 16, 1955.

L. A. Finzi and G. C. Feth are with the Carnegie Institute of Technology, Pittsburgh, Pa.

This work has been supported in part by the U.S. Office of Naval Research.

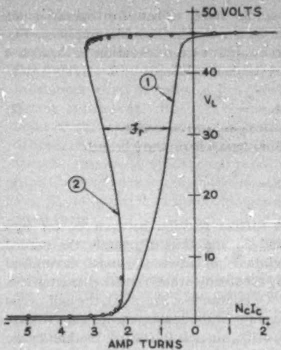


Fig. 1. Transfer characteristics

Curve 1—Experimental characteristic of stable self-saturating center-tap amplifier with resistive load

Curve 2—Calculated bistable characteristic of the same amplifier modified by overregeneration. Circled points are experimental checks

l, 74
1
55

depends to various extents on core geometry and material, on a-c gate frequency and voltage per turn, on the load, and also on the suppression index (i.e., on the impedances of the paths offered to the flow of even harmonic currents). Knowledge of how this loop is described in time yields the instantaneous mmf $Hl=\Sigma Ni$ of one core and thus also the half-cyclic average ΣNI, where the summation is extended to all windings on the core as per Ampere's circuital law.

Contributions to the total core mmf are given by signal currents, bias currents, currents of additional feedback windings, and gate currents (which may include effects of rectifier leakage and also circulating currents forced by load circuit inductances). For a chosen α, modifications of any of these components of the total core mmf can be compensated by proper changes of the others. (This applies under the restriction that the variables which determine the minor loop and the way it is described in time, are not changed significantly in the process. For instance, the total suppression index should not be changed by the introduction of low impedance bias windings in an otherwise highly constrained situation; the a-c volts per turn should be properly adjusted when crossover booster windings are used; etc.)

All this is well recognized, e.g., in the common techniques used to predict transfer characteristics of saturable reactors modified by the addition of external feedback. The prediction of the transfer characteristic of an amplifier made bistable by overregenerative mmf is somewhat more critical and hardly can be based on the results of the oversimplified analyses commonly used to describe the transfer characteristic of the simple saturable reactor. A more realistic starting point is offered by transfer characteristics of self-saturating amplifiers, either calculated carefully under consideration of the core and rectifier properties and loading[3] or, even better, measured experimentally.

The procedure is reviewed in the example of Fig. 1. Curve 1 is the experimental characteristic of a self-saturating center-tap amplifier, to be described later in further detail, with a given resistive load, a-c voltage, and frequency, presented in terms of average-control ampere turns N_cI_c and of gross output voltage V_L. V_L is the half-cyclic average of that part of the gate voltage which is not absorbed by the cores; i.e., $V_L \cong RI_g$ where I_g is the half-cyclic average of gate currents and R is the total resistance of load, gate winding on one core and rectifier in forward conduction. $R=47.3$ ohms in this example. Curve 2 is a bistable characteristic calculated from curve 1 after introduction of an overregenerative mmf of feedback $\mathfrak{F}_r=N_rI_g$, (in this instance $N_r=3$ turns per core). The mmf of control for any given output V_L is predicted by adjusting the corresponding control mmf of curve 1 by the amount needed to compensate for the additional mmf $\mathfrak{F}_r$ introduced. The circled points confirm experimentally the predictions of the calculated curve 2 over its stable regions.

Bistable characteristics are obtained in practice also in other manners. Feedback of load currents, rather than gate currents, yield somewhat different results if the load is rectified-inductive with relaxation paths. Cross-connected boosting coils, each connected in series with the gate coil on the other core, have been suggested[4] and discussed.[5] More frequently the mmf $\mathfrak{F}_r$ is obtained by currents proportional to the output voltage V_L, thus making the bistable transfer characteristic somewhat independent of changes of load. The mmf $\mathfrak{F}_r$ conceivably can be made proportional to the half-cyclic voltage absorbed by either core in the prefiring intervals; then the transfer characteristic of the bistable amplifier nearly coincides with the transfer characteristic of the original stable amplifier in the region of large outputs, departing from it and bending more and more to the right side of Fig. 1 for decreasing outputs. Pseudoregenerative effects of rectified inductive loading without relaxation paths are also utilized sometimes to obtain a bistable characteristic.

Two or more of these methods of feedbacking may be used jointly or in mutual opposition to satisfy certain re-

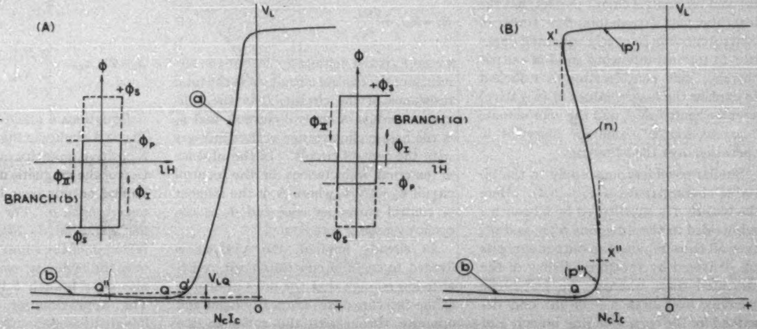


Fig. 2. Branches of stable transfer characteristic and of bistable characteristic

A—Branches a and b of stable transfer characteristic

B—Branches s and b of bistable characteristic

Branch s consists of region p', n, and p'' quirements of the designer, e.g., to improve on the "sharpness" of either corner of the characteristic. (Poorly defined corners yield poorly resolved "switching" action and imply also slower response, as will be shown.) In any case, the basic concept reviewed in the example of Fig. 1 leads to pertinent predictions.

Branches of Transfer Characteristics

Fig. 2(A) shows the transfer characteristic of a stable magnetic amplifier, i.e., the characteristic of a self-saturating amplifier for given loading, suppression index, gate frequency, and voltage, in conditions of nearly "normal" a-c supply voltage. For any assigned control mmf, $N_c I_c$, the output voltage V_L (half-cyclic average or also rms) is determined uniquely. On the other hand, a wanted V_L can be obtained by the use of two values of control $N_c I_c$, one on the branch *a* of the characteristic and the other on the branch *b*.

The phenomena taking place over branch *b* differ somewhat from those of branch *a*. Over branch *a* the core I to which line voltage is applied in the conducting direction of gate rectifier in the prefiring interval of the half-cycle considered (e.g., positive half-cycle) rises from some initial preset flux value Φ_p towards saturation $+\Phi_s$ (while core II resets from $+\Phi_s$); thus firing occurs and an output current is delivered to the load through the gate winding of the saturated core. Over branch *b* in the prefiring interval of the positive half-cycle, core I ascends from $-\Phi_s$ while core II descends from some initial value Φ_p' towards $-\Phi_s$. At some time core II reaches negative saturation and firing occurs; in this case, however, the output current is delivered by the gate winding of the unsaturated core I. The transformer coupling that exists then between an output gate winding and the control circuit implies that the instantaneous magnetomotive force of control currents be nearly equal and opposite to the instantaneous mmf of output current; this consideration is sufficient to explain the larger values of (negative) average control $N_c I_c$ and the approximate "current source" behavior observed in operation over the *b* branch.

Similar considerations apply to the bistable characteristic of Fig. 2(B). Here the branch *a* is substituted by a branch *s* subdivided in three regions *p'*, *n*, and *p"*; over all three regions the output currents are delivered by the gate winding of the saturated core, whereas the half-cyclic behavior over branch *b* is the one described for the corresponding branch *b* of Fig. 2(A). Several values of output V_L are possible for a given control $N_c I_c$. The region *n* is one of unstable operation, as discussed later, yet its points are actually traversed in the transition from low to high output or vice versa. This can be seen qualitatively in the experimentation by simultaneous observation of meters in the control and output circuits and is evidenced by the oscillographic transfer characteristic of Fig. 3 (obtained under use of half-cyclic averaging networks at the oscilloscope terminals to reduce blurring of the trace).

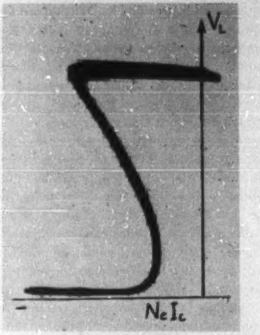


Fig. 3. Locus of points V_L, $N_c I_c$ described in switching process

Mechanism of Transition

Transitions of the bistable amplifier from the one output level to the other are treated by extending the approximate approach accepted for the "slow-transient" analysis of conventional amplifiers with single-valued output. The voltage differential equation of the control circuit is written

$$E_c = R_c I_c + \frac{d\psi_c}{dt} \qquad (1)$$

where E_c is the voltage applied at the terminals of the control circuit, R_c is the total resistance of the circuit, I_c is the half-cyclic average of control current, and ψ_c is the half-cyclic average of flux linkages with the control circuit. In the absence of external inductances in the control circuit $\psi_c = 2N_c\Phi_0$ where N_c is the number of control turns per core and Φ_0 is the cyclic average of core flux.

As already implied, the analysis is limited to cases of control circuit resistance low enough that the rates of change of the two core fluxes are nearly equal and opposite throughout the cycle in the steady state. Then for operation on branch *s* of the bistable amplifier, Fig. 2(B), just as for operation on branch *a* of stable amplifiers

$$\Phi_0 = \frac{\Phi_s + \Phi_p}{2} \qquad (2)$$

For operation on the *b* branch

$$\Phi_0 = \frac{-\Phi_s + \Phi_p'}{2} \qquad (3)$$

A further relation between Φ_p (or Φ_p') and I_c, and thus ultimately the needed relationship between ψ_c and I_c, is supplied by the steady-state transfer characteristic.

Over branch *s*, Fig. 2(B), the half-cyclic average of the instantaneous gate voltage $v_g = V_{gm} \sin \omega t$ is $2/\pi\ V_{gm}$; the half-cyclic average V_{cores} of the voltage absorbed by the cores is

$$V_{\text{cores}} = \frac{\omega}{\pi}(\Phi_s - \Phi_p)N_g \qquad (4)$$

where N_g is the number of gate winding turns per core. The difference $(2/\pi V_{gm} - V_{\text{cores}})$ is the half-cyclic average V_L of gross output voltage. Under introduction of the concept of "normal" core voltage crest $V_{gN} = \omega N_g \Phi_s$ this becomes

$$\begin{aligned} V_L &= \frac{2}{\pi}V_{gm} - \frac{\omega}{\pi}(\Phi_s - \Phi_p)N_g \\ &= \frac{2}{\pi}V_{gm} - \frac{1}{\pi}V_{gN} + \frac{\omega}{\pi}\Phi_p N_g \qquad (5) \end{aligned}$$

whence, under use of equation 2

$$\psi_c = 2N_c\Phi_0 = \frac{2}{\omega}\frac{N_c}{N_g}\left[(V_{gN} - V_{gm}) + \frac{\pi}{2}V_L\right] \qquad (6)$$

Over branch *b*, Fig. 2(B) or 2(A), these equations modify as follows

$$V_{\text{cores}} = \frac{\omega}{\pi}(\Phi_p' - (-\Phi_s))N_g \qquad (7)$$

$$\begin{aligned} V_L &= \frac{2}{\pi}V_{gm} - \frac{\omega}{\pi}(\Phi_s + \Phi_p')N_g \\ &= \frac{2}{\pi}V_{gm} - \frac{1}{\pi}V_{gN} - \frac{\omega}{\pi}\Phi_p' N_g \qquad (8) \end{aligned}$$

$$\psi_c = 2N_c\Phi_0 = -\frac{2}{\omega}\frac{N_c}{N_g}\left[(V_{gN} - V_{gm}) + \frac{\pi}{2}V_L\right] \qquad (9)$$

Equations 6 and 9 in conjunction with Fig. 2(A) indicate that the term $d\psi_c/dI_c$ is a positive quantity at all points of operation of the magnetic amplifier with single-valued output over branch *a* as well as over branch *b*. For the bistable amplifier, Fig. 2(B), $d\psi_c/dI_c$ is positive over the region *p'* of the *s* branch but it is negative over the region *n;* over the region *p"* and also over branch *b* it is again positive. The approximation $E_c = 0 = R_c$ underlying these derivations may be objectionable in its use over parts of the *b* branch far to the left, in view of the comparatively large signal voltages and large voltage drops in the control circuit resistances, even in cases of low-control circuit resistances. But the operating point of the amplifier at its stable level of lower output is on an initial part of this branch; expression 9, although resulting from a simplified reasoning, indicates that $d\psi_c/dI_c$ is positive and comparatively small there.

It may be added that for conditions of underexcitation, $V_{gm} < V_{gN}$ or, more precisely, $V_{gm} < V_{gN} + \pi/2\ V_{LQ}$ where V_{LQ} is the half-cyclic average of the voltage drop of quiescent current through the gate circuit resistance, a singular point Q, or a small region thereabout appears between the branches *a* and *b* of the characteristic of the stable amplifier, Fig. 2(A) and between branches *s* and *b* of the characteristic of the bistable amplifier. At that point, saturation of the cores does not occur in the cycle; both core fluxes swing up and down about some average value Φ_0 which cannot be evaluated immediately from the kind of information supplied by the idealized transfer characteristics shown. That is, ψ_c may have any intermediate value between the two given by equations 6 and 9 at $V_L = V_{LQ}$.

On the other hand, for conditions of overexcitation $V_{gm} > V_{gN}$ or, more precisely, $V_{gm} > V_{gN} + \pi/2\ V_{LQ}$, a horizontal segment $Q'Q''$ of minimum output appears, separating branches *a* and *b* as shown by dashed line in Fig. 2(A), or separating the branches *s* and *b* of the bistable characteristic, not shown. For operating points over this horizontal line both cores, if well matched, seem to swing from Φ_s to $-\Phi_s$ and vice versa during a cycle; hence one can say that $\psi_c = 0$.

It is well known to the experimenter that the region of lowest outputs of any transfer characteristic cannot be obtained precisely and that it is difficult to define the point Q of minimum output. Also, more complex and as yet poorly understood phenomena, e.g., triggering and effects of poor core matching, may intervene in this region. The foregoing remarks attempt to describe more completely the whole transition of the bistable amplifier from the one stable output level to the other. The times involved to pass through these debatable points of the characteristics seem not very significant in the experimentation; the considerations just given will be used further for a semiqualitative justification of this finding.

Equation 6 indicates that $d\psi_c/dI_c$ is a nearly constant quantity over the "linear range" of the transfer characteristics of conventional amplifiers. In these cases equation 1 reduces to equation 10

$$E_c = R_cI_c + L_c dI_c/dt \tag{10}$$

where $L_c = d\psi_c/dI_c$ is the "apparent differential inductance of the control windings," which is assumed constant. This linear differential equation with constant coefficients is the basis for many familiar treatments used to derive expressions for time constants and figures of merit.

A similar linearization may be used tentatively over the *n* region of the characteristic of the bistable amplifier. Notice, however, that L_c becomes a negative number. This recognition is sufficient to explain the unstable behavior of this amplifier at any operating point over the region *n*. On the other hand, this linearization is less adequate to describe with quantitative accuracy the switching action of the amplifier from the one output level to the other. In a build-down from high to low output, the quantity $d\psi_c/dI_c$, initially positive and rather small over the portion p', increases gradually to quite large values before becoming $+\infty$ at the limit point of stability X'; it jumps there to $-\infty$ and acquires then finite and rather large negative values over the *n* region. At the limit point of stability X'' $d\psi_c/dI_c$ becomes again $-\infty$ and jumps there to $+\infty$, acquiring finally finite positive values over the regions p'' and *b*. The same processes take place, in the reverse order, in a build-up from low to high output level. A large amount of the total switching time is spent passing through the region of the characteristic around the first limit of stability met, X' or X'' as the case may be. Predictions of over-all time response based on the linearized approach of equation 10 are inadequate and one must undertake the solving of equation 1 by graphical or numerical procedures.

Consider, for instance, in Fig. 4 an amplifier initially at a saturated output level V_{Li} and let a step of signal be applied so that the control voltage becomes $E_c = R_cI_{cf}$.

For the *s* branch equation 1 is rewritten in terms of the time-varying half-cyclic terminal quantities I_c and V_L under use of equation 6 as

$$E_c - R_cI_c = E_a = \frac{\pi}{\omega}\frac{N_c}{N_g}\frac{dV_L}{dt} \tag{11}$$

The voltage $E_a = E_c - R_cI_c$ shall be called the "actuating voltage;" the rate of

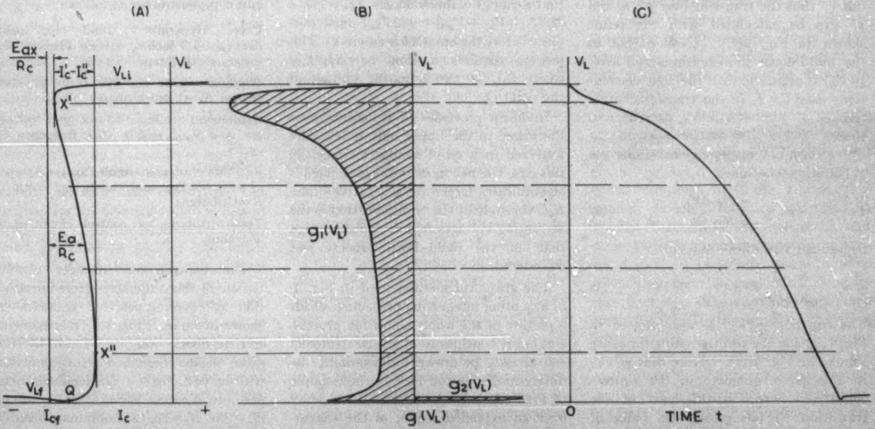


Fig. 4. Graphical integration of equation 1 to describe output V_L as function of time in build-down transition

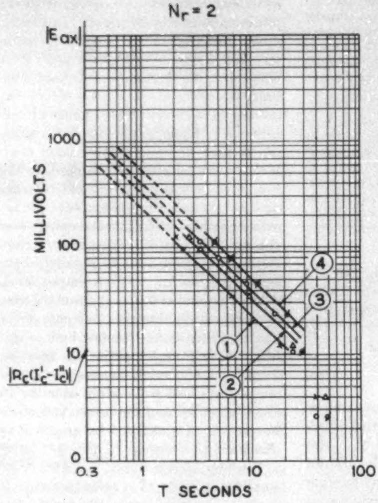


Fig. 5 (left). Experimental investigation of influence of E_{ax} upon response. Overregenerative current feedback. $N_r = 2$

Build-down process:
Curve 1—Time to 63.2% of total swing
Curve 2—Time to 86.5% of total swing

Build-up process:
Curve 3—Time to 63.2% of total swing
Curve 4—Time to 86.5% of total, swing

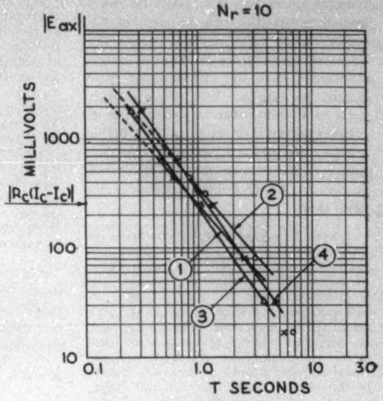


Fig. 6 (right). Same as Fig. 5 with $N_r = 10$

change of the output V_L at any point of the characteristic traversed is proportional to the corresponding E_a.

By separation of variables

$$dt = \frac{\pi}{\omega}\frac{N_c}{N_g}\frac{1}{E_a}dV_L = g_1(V_L)dV_L \qquad (12)$$

The function $g_1(V_L)$ is plotted versus V_L in Fig. 4(B). According to equation 12 the shaded area between any two values of V_L is proportional to the time required for the change from the one value to the other; thus the transient over p', n, and p'' can be calculated with the result shown in Fig. 4(C). (Notice that in the build-down process considered here, $g_1(V_L)$ is negative, with the sign conventions used for I_c in the transfer characteristic of Fig. 4(A); V_L decreases in time.)

For the b branch, equation 1 under use of equation 9 becomes

$$E_c - R_cI_c = E_a = -\frac{\pi}{\omega}\frac{N_c}{N_g}\frac{dV_L}{dt} \qquad (13)$$

Separating variables results in

$$dt = -\frac{\pi}{\omega}\frac{N_c}{N_g}\frac{1}{E_a}\,dV_L = g_2(V_L)dV_L \qquad (14)$$

The function $g_2(V_L)$ is also plotted in Fig. 4(B) and the corresponding times are shown in Fig. 4(C). (Note that $g_2(V_L)$ is a positive quantity in the process considered here; in this part of the transition V_L has passed the value of minimum output and is rising over the b branch towards the final value V_{Lf}.)

The transfer characteristic considered was based on a situation of normal excitation. If overexcitation is used, the branches s and b are joined by a horizontal segment which is traversed in nearly no time, as stated in the paragraphs following equation 9. If, on the other hand, underexcitation is used, an ambiguity of values of ψ_c occurs at the point Q of transition from s to b. By comparison of the values of ψ_c given by equations 6 and 9 for $V_L = V_{LQ}$, it appears that a change of flux linkages $\Delta\psi_c = -4/\omega\; N_c/N_g[(V_{gN} - V_{gm}) + \pi/2V_{LQ}]$ must take place before the branch b is entered. This process requires a time $\Delta t = \Delta\psi_c/E_{aQ}$, where E_{aQ} is the actuating voltage at the point Q.

Build-up processes of transition from the lower to the upper output level are analyzed in a quite similar manner; in this case the net signal voltage applied is algebraically larger (less negative) than R_cI_c throughout the transient, that is the activating voltages E_a have opposite sign than in the build-down process just considered.

The graphical method used in Fig. 4, or any other integrating procedure, yields a picture of the whole transition process. But many conclusions of practical interest can already be drawn by examining the differential equation 11. In the example of Fig. 4 it is evident that $|E_a|$ is small over an extended region of the characteristic about the point X'. The corresponding time required will be a large part of the whole. This justifies the large contribution given by this region to the total shaded area of Fig. 4(B). As this region is passed, $|E_a|$ becomes considerably larger and correspondingly the remaining part of the process is faster. Similarly in a build-up process E_a is smallest over the region about X''. The passage over this region is most time-consuming while the remaining part of the process is faster.

Experimentation

Extensive experimentation has been performed on a typical bistable amplifier in center-tap (d-c output) connection. The data of the amplifier and related circuit components are as follows:

Cores: Hypernik V, 2-mil tape, inside diameter 1.5 inches, outside diameter 2.5 inches, width of tape 0.5 inches

Windings: Gate $N_g = 993$ turns per core, control $N_c = 1{,}240$ turns per core

Germanium rectifiers *IN158*, gate voltage per core $V_{g\,\mathrm{rms}} = 52.5$ volts, frequency 60 cycles

Resistances of gate circuit (load, resistance of one winding and conducting rectifier) $R = 47.3$ ohms

Total resistance of control circuit $R_c = 36.5$ ohms

The experimental transfer characteristic of this amplifier, before introduction of overregenerative feedback is shown in curve 1, Fig. 1. Overregenerative feedback was introduced by the same means described in conjunction with curve 2, Fig. 1. The feedback turns per core N_r were varied over the range $N_r = 2$ to $N_r = 10$, thus obtaining bistable characteristics with a variety of widths to the loop of instability.

Calculations based on the graphical integration method of Fig. 4 were compared with times of response measured with the oscillographic techniques described by Critchlow in reference 6. (Simple observations of envelopes of peaks would have been particularly meaningless in that part of the transient for which the angle of firing varies between 0 and 90 degrees.) The agreement which was found under the experimental conditions as stated was extremely good throughout.

In view of the relative importance of the time of passage around the first point of instability it appeared of interest to investigate whether some simple empirical relation could be established between the actuating voltage $|E_{ax}|$ at that point and the time interval T required to reach a certain percentage (e.g., 63.2 per cent (%) or 86.5%) of the over-all swing of V_L in transitions from saturated to quiescent output or vice versa.

Figs. 5 and 6 present typical results obtained for the two extreme cases $N_r=2$ and $N_r=10$. It appears that for a large range of values of $|E_{ax}|$, a general relation of the type $T=K/|E_{ax}|$ prevails. This is the kind of relation that would be predicted for an actuating voltage E_a constant throughout the switching process, (as would be the case for a vertical transfer characteristic upon application of a constant control voltage E_c); it differs noticeably from the relation that would result from a linearization of the bistable characteristic into the shape of an inverted Z. The relative importance of the passage through the region around the first corner is evidenced once more; over this sluggish region E_a is more or less constant, at least as long as $|E_{ax}|$ is not too small. The relation of inverse proportionality stated applies also for larger values of $|E_{ax}|$ than shown by the experimental points plotted in Figs. 5 and 6, (the extension by dotted lines is somewhat justified by the results of tests performed with other values of N_r, not shown). Finally some explanation can be given for the different values of K seen in Fig. 5, e.g., for the build-up and build-down processes to 63.2%; the upper corner of the pertinent bistable characteristic had smaller radii of curvature than the lower corner. The constants K decrease as the width of the loop is broadened. This is so for two reasons; wider loops have inherently sharper corners, i.e., the critical region to be passed is less extended; moreover, as this region is passed, the actuating voltages are larger, thus the remaining part of the process is also faster.

Additional Remarks

Equations 11 and 12 indicate that proportionally faster response is obtained for a given bistable amplifier if the signal voltage E_c and the control resistance R_c are increased both at the same rate. The concept of a "figure of merit" can be introduced for any given E_c/R_c ratio. This concept is not very meaningful, however, because the change of power output is a fixed quantity for given loading and because even better improvements on the response time can be obtained if one increases E_c within certain limits without varying R_c. Some thought must be given then to the point V_{Lf} of steady-state output that results on the characteristic (even though the b and p' regions are remarkably flat) and on the current-carrying capacity of the control winding.

The treatment given is based on a general simplified reasoning which applies not only to the bistable amplifier but to any amplifier with low-suppression index. (This treatment has been repeated here, perhaps in more detail than usual, because many of the points raised cannot be ignored or by-passed with the simpler criteria that are common when dealing with more conventional stable amplifiers operating over the linear part of their transfer characteristics.) With this recognition certain familiar statements can be extended to the bistable amplifier without need for any renewed proof.

Introduction of an mmf of bias $\mathfrak{F}_b=N_bI_b$ simply shifts the origin of the abscissas of Fig. 2(B) by an amount $\pm\mathfrak{F}_b$ as the case may be. The transient analysis, developed in the foregoing under the assumption that no other circuit existed directly coupled with the control circuit, is modified only to the extent that the right-hand term of equations 11 and 12 is multiplied by the factor $(N_c^2/R_c+N_b^2/R_b)/N_c^2/R_c$ where R_b is the total resistance of the bias circuit. A further circuit appears coupled in the build-down process of bistable amplifiers of the "doubler" type, because significant circulating currents in this case can be induced to flow in the loop formed by the two gate windings. For this situation a term $N_g^2/2r_g$ is introduced in the summation of the turns square per ohm in the foregoing, where r_g is the resistance of one gate winding and associated rectifier in forward conduction.

The assumption of zero-suppression index here is subject to the same limitations as in more conventional cases. By increasing R_c the bistable characteristic changes; however, the procedure which is used in Fig. 1 can be applied reliably again to obtain the pertinent bistable characteristic if one starts from a stable characteristic measured with the value of R_c in question. The accuracy of the predictions of transient response decreases with increasing suppression index. Calculations with the methods of Fig. 4 and measurements with the techniques of reference 6 were found to agree very well over a wide range for the experimental conditions described (suppression index $\sigma=0.00125$); but the calculations of $T_{63.2\%}$ were shorter than measured, by about 25% at $\sigma=0.125$; this error became twice as great at the higher suppression $\sigma=0.625$. (The analysis given falls short at high R_c because the time average of flux Φ_0 is not properly given by the arithmetic mean expressed in equations 2 and 3 and because overlap is more pronounced. It is beyond the modest scope of this paper to open an investigation of the general behavior of magnetic amplifiers with even-harmonics suppression.)

Finally, the question of the response of the bistable amplifier to time-varying signals deserves some brief comments. If E_c varies during the transition process, step-by-step methods are in order for the integration of equations 11 and 12 so as to use proper values of E_c for every value of V_L. A particular case of interest occurs when one wishes to produce a switching of output levels by means of a change of signal applied in the form of a pulse of limited duration. The question arises as to how long this duration must be. If the applied change of signal voltage is removed before steady state is reached the process may either proceed in the initial direction or it may reverse. A reversal of the process occurs if the sign of E_a (and therefore the sign of dV_L/dt) changes when the signal is changed.

A similar problem arises for bistable amplifiers, designed with a wide loop of instability and provided with self-biasing circuits intended to reduce the swing of signal required to cause a switching action. A self-biasing of this kind is readily obtained, for instance, if the load of the bistable amplifier is another switching device, e.g., an electromagnetic relay or another bistable amplifier. This device, once it is closed, supplies an auxiliary mmf which biases the amplifier preparing it for a switch-off action for small changes of E_c. But sustained oscillations of the system are likely to occur if that switching device closes already at some intermediate value of the ascending V_L; the consideration just given yields the proper amount of self-bias allowable.

Summary

Widely accepted approaches that underlie the study of the transient response of common amplifiers with low-suppression indexes have been applied to describe the behavior of bistable amplifiers in the transition from one stable output level to the other.

Certain issues that can be ignored or by-passed for cases of stable amplifiers operating over the linear range of their transfer characteristic have received a renewed and more detailed treatment, as needed. The influence of various pertinent factors upon the mechanism of response has been discussed and illustrated experimentally. Limitations inherent to the method of analysis chosen have been mentioned.

References

1. Magnetic Amplifiers (book), H. F. Storm. John Wiley & Sons, New York, N. Y., 1955, p. 200.

2. *Ibid.*, pp. 84, 309.

3. The Operation of Magnetic Amplifiers with Various Types of Loads, Part II—Controlling the Angle of Firing, L. A. Finzi, R. R. Jackson. *AIEE Transactions*, vol. 73, pt. I July 1954, pp. 279–288.

4. The Transductor Amplifier (book), V. Krabbe. Lindska Bikteyskerict, Orebo, Sweden, 1947, p. 69.

5. The Effective Feedback Ratio of Magnetic Amplifiers, L. A. Finzi, G. F. Pittman, Jr., H. L. Durand. *AIEE Transactions*, vol. 71, pt. I, Apr. 1952, pp. 157–64.

6. Oscillographic Techniques for the Evaluation of Magnetic Amplifier Response, D. L. Critchlow. *Ibid.*, vol. 74, pt. I, Nov. 1955, pp. 607–10.

Standards for the Presentation of Data on Magnetic Amplifier Core Materials

AIEE COMMITTEE REPORT

THE Materials Subcommittee of the Magnetic Amplifier Committee was organized principally to set up standards on materials used in connection with magnetic amplifiers and to encourage the interchange of information on materials. The object of one of the first projects which the Subcommittee undertook was to develop a set of standards to guide manufacturers in presenting basic data on magnetic amplifier core materials. These standards have been revised somewhat since their inception in 1951 and were formally submitted to the Magnetic Amplifier Committee for adoption in February 1955.

It is the purpose of this report to present these standards as adopted by the Magnetic Amplifier Committee and to give a brief progress report of the Materials Subcommittee's activities.

Some emphasis should be placed on the fact that the standards given here are intended for use in presenting basic information on core materials and are not proposed as a test method for grading and matching magnetic amplifier cores. The Materials Subcommittee has recognized the need for the study and standardization of methods of grading and matching cores and as a result has organized a Working Group on Core Matching and Grading to study the various methods now being used by manufacturers of core materials and magnetic amplifiers. In general, matching and grading techniques are designed to give data in production which correlates with actual magnetic amplifier performance. From a survey of a large number of core manufacturers and consumers it was found that the majority of core matching and grading methods employ sinusoidal current rather than sinusoidal voltage excitation because of the greater simplicity of the sinusoidal current test and its adaptability to production testing. This working group is now attempting to set up standards for the three most popular types of grading and matching methods.

The Materials Subcommittee has also organized a Working Group on Constant Current Reset to study other tests which may be classified as basic material tests and which may become a substitute for or a supplement to the standards given herein. The working group is considering the constant-current flux reset and the half wave self-saturating magnetic amplifier circuits as probable future basic core material tests.

Another Working Group on Standardization of Toroidal Core Sizes has been formed to attempt to standardize tape-wound toroidal core sizes. This standardization, when complete, will enable the designer to select cores which are readily available and will assist the manufacturer of cores in the standardization of tools and fixtures.

The Materials Subcommittee has been carrying on close liaison with the newly formed Magnetic Amplifier Rectifier Subcommittee of the Metallic Rectifier Committee by having several members common to both subcommittees. This Subcommittee is currently working on definitions and test circuits pertaining to magnetic amplifier rectifiers.

The foregoing is a brief summary of the past and present activities of the Materials Subcommittee and points out some of the most important projects which will be continued or originated in the future. The standards for presenting data on magnetic amplifier core materials are given in the following. It should be pointed out that the standard test core sizes listed are not necessarily proposed as standard magnetic amplifier core sizes but are recommended for use when performing the tests described here so that the variable of physical core proportions will be a constant when comparing data from several manufacturers.

Magnetic Data

D-C Magnetization Curve

1. To $\mathbf{H}=10$ oersteds for cases where the intrinsic flux density increases 5 per cent or less between $\mathbf{H}=10$ and $\mathbf{H}=50$ oersteds.
2. To $\mathbf{H}=50$ oersteds for cases where the intrinsic flux density increases more than 5 per cent between $\mathbf{H}=10$ and $\mathbf{H}=50$ oersteds.

D-C Major Hysteresis Loop

1. To a peak magnetizing force of $\mathbf{H}=1$ oersted if the d-c coercive force is less than 0.05 oersted.
2. To a peak magnetizing force of $\mathbf{H}=10$ oersteds if the d-c coercive force is 0.05 oersted or greater.

Major Dynamic Hysteresis Loops

1. Standard frequencies
   a. 60,400, 1,000, and 10,000 cycles per second as applicable.

Paper 55-540, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 30, 1955; made available for printing May 3, 1955.

The personnel of the AIEE Materials Subcommittee of the Committee on Magnetic Amplifiers are: P. L. Schmidt, *Chairman*, A. O. Black, W. C. Bullock, J. R. Conrath, E. A. Gaugler, H. A. Goldsmith, E. A. Harty, E. W. Lehtonen, F. J. Lingel, M. F. Littman, H. W. Lord, V. J. Louden, J. E. Mitch, G. F. Pittman, W. T. Sackett, W. F. Schmidt, D. J. Scorgie, D. J. Sikorra, and E. J. Smith.

Vol
Pt.
195

b. In cases where the manufacturer desires to test at other frequencies, the following are preferred: 2 kc, 5 kc, 100 kc, 500 kc, 1 megacycle, and 10 megacycles.

2. Excitation

a. The induced voltage shall be sinusoidal with less than 10-per-cent total harmonic distortion.

b. The peak magnetizing forces shall be 2 times the a-c coercive force value and 20 times the a-c coercive force value or 10 oersteds, whichever is smaller for each of the frequencies at which tests are made.

Standard Test Core Sizes

1. Toroidal Core—1$^5/_8$ inches inside diameter by 2 inches outside diameter by 1/4 inch high for frequencies up to 1,000 cycles per second.

2. Toroidal Core—1/2 inch inside diameter by 3/4 inch outside diameter by 1/8 inch high for frequencies of 1,000 cycles per second and greater.

3. A core composed of double-lapped U-punchings having a cross-sectional area of 1/4 inch (leg) by 1/2 inch (stack height) and a mean length approximately equivalent to the larger of the two toroidal cores.

Temperature

1. The tests shall be conducted at a core temperature of 25±10 degrees centigrade.

2. In cases where the manufacturer desires to test at other temperatures, nominal core temperatures of −55, 105, 150, and 200 degrees centigrade are preferred.

Influence of ID/OD Ratio on Static and Dynamic Magnetic Properties of Toroidal Cores

R. W. ROBERTS
ASSOCIATE MEMBER AIEE

R. I. VAN NICE
ASSOCIATE MEMBER AIEE

Synopsis: The optimization of toroidal core shapes for use in self-saturating magnetic amplifiers requires a knowledge of the influence of the inside-diameter/outside-diameter (ID/OD) ratio on apparent core magnetic properties. A theoretical result has been derived assuming a parallelogram-shaped basic hysteresis loop. Since this result has several unexpected features of great practical interest, a rather elaborate experiment was performed which verified their existence. In one application, it has been possible to double the gain of a push-pull amplifier stage by changing the core height and radial build-up so as to increase the ID/OD ratio without changing the iron area.

THE magnetic properties of a closed core in which all flux lines do not have nearly the same length differ from those of a thin ring sample of the same material. Although this phenomenon is not of major importance in transformer design, it does appreciably influence the characteristics of high-performance self-saturating magnetic amplifiers. The need for an investigation of the effect of the ID/OD ratio of toroidal cores on their magnetic properties arises when attempts are made to calculate optimum core shapes for magnetic amplifier applications. Satisfactory expressions are known for the effect of the other core shape factors affecting amplifier performance, e.g., mean magnetic path length, cross-section area, and window area. A simple experiment using two cores indicated that the ID/OD effect might be quite large, but no satisfactory expression for it was found during a review of the literature.

McConnell[1] has investigated the influence of the ID/OD ratio on static properties of toroidal cores, but his results are inconclusive. Dieterly[2] has published a study of dynamic hysteresis loops under sine flux conditions showing that loops for cores with low values of ID/OD ratio are sheared over with respect to the loop for the basic material. The application of Dieterly's results to magnetic amplifier performance is not apparent because of the lack of any known correlation between dynamic loops and the performance of amplifiers with relatively high control circuit impedance.[3]

The approach taken in the present paper is as follows: If a relationship could be found between amplifier performance and a hypothetical (effective) dynamic hysteresis loop, then an effective loop could be determined experimentally for a thin core, i.e., one with ID/OD ratio approximately unity. Such a loop expresses a basic property of the magnetic material since it is not influenced by core geometry. The effective hysteresis loop corresponding to any ID/OD ratio might then be calculated using an approach similar to that of McConnell. It is then to be hoped that the previously mentioned relationship between amplifier performance and effective hysteresis loop could be used successfully to predict amplifier performance for any ID/OD ratio.

Lehmann[3] and others[4,5] have pointed out the relation between amplifier performance and a **B-H** plot called a control magnetization curve. The latter, although actually a locus of points on a series of minor hysteresis loops, has a form similar to a portion of the left flank of a major hysteresis loop. Carrying out this analogy by adding a symmetrical right flank, an effective hysteresis loop corresponding to magnetic amplifier operation is obtained, and the program of the previous paragraph becomes possible to carry out.

At the outset the work reported here was restricted to toroidal cores, though similar results could probably be obtained for other shapes. Reasons for this restriction included the mathematical and experimental simplicity of the toroidal geometry as well as the practical importance of toroidal cores in magnetic amplifier applications.

Included in this paper are theoretical treatments of the influence of the ID/OD ratio on the hysteresis loop for several different types of basic material loops, and the results of an experimental investigation of the influence of the ID/OD ratio on the static loop and a-c control magnetization curves for a typical magnetic amplifier core material. The results are compared to determine the usefulness of the theoretical treatment.

Derivation of Equations

A theoretical calculation will now be made of the dependence of the B-H relation of a toroidal core on the ID/OD ratio. Several physically reasonable simplifying

---

Paper 55-538, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 29, 1955; made available for printing May 12, 1955.

R. W. ROBERTS and R. I. VAN NICE are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

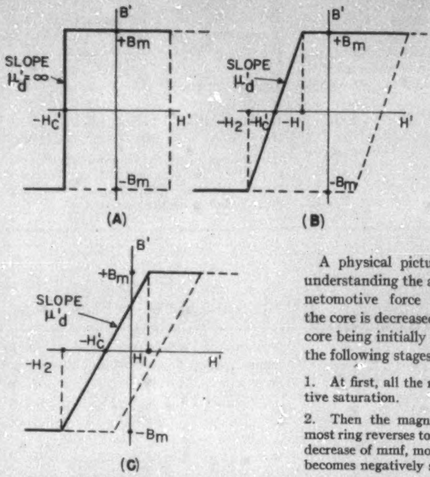


Fig. 1. Assumed forms of B-H loops

assumptions have been found necessary. The basic assumption is that a tape-wound toroidal core can be replaced by an equivalent core consisting of a large number (infinite in the limit) of thin closed concentric rings. The need for this assumption is essentially geometrical, since a helical core with insulating gaps between wraps is difficult to describe mathematically.

A second assumption is that a thin ring of the magnetic material has a parallelogram-shaped **B**-**H** loop. Three different cases will be considered: 1. perfectly rectangular loop, Fig. 1(A); 2. loop with sloping flanks, Fig. 1(B); and 3. loop with remanent flux density $\mathbf{B}_r$ less than maximum $\mathbf{B}_m$, Fig. 1(C). While case 1 is derivable as a special instance of case 2, its simplicity suggests its independent derivation as an introductory example. The meaning of case 1 in terms of magnetic theory is that all the magnetic dipoles have the same coercive force; thus the slope of the saturated region is zero, and the slope of the flanks is infinite. The coercive forces of the dipoles in cases 2 and 3 have a rectangular distribution about a mean value.

CASE I. PERFECTLY RECTANGULAR LOOP

It is desired to find the average flux density **B** of a toroidal core with inner and outer radii R_1 and R_2 respectively as a function of the magnetic field intensity **H** at the mean radius $1/2(R_2+R_1)$. The method to be used is identical with that of McConnell.[1]

A physical picture will be helpful in understanding the analysis. As the magnetomotive force (mmf) impressed on the core is decreased slowly from zero, the core being initially at positive saturation, the following stages can be visualized:

1. At first, all the rings are at $+\mathbf{B}_m$, positive saturation.
2. Then the magnetization of the innermost ring reverses to $-\mathbf{B}_m$, and with further decrease of mmf, more and more of the core becomes negatively saturated.
3. Finally, the magnetization of the entire core is at $-\mathbf{B}_m$.

The average flux density may be calculated from

$$\mathbf{B}=\frac{1}{R_2-R_1}\int_{R_1}^{R_2}\mathbf{B}'dR \qquad (1)$$

where $\mathbf{B}'$ is the flux density in the ring of radius R and thickness dR. In this case, $\mathbf{B}'$ takes on only two values

$$\begin{aligned}\mathbf{B}'&=-\mathbf{B}_m,\ R_1<R<R_x\\&=+\mathbf{B}_m,\ R_x<R<R_2\end{aligned} \qquad (2)$$

where R_x is the radius of the surface which separates the two oppositely magnetized regions of the core. The field intensity in a ring $\mathbf{H}'$, as a function of mmf NI, is

$$\mathbf{H}'=\frac{0.4\pi NI}{2\pi R}=0.2\frac{NI}{R} \qquad (3)$$

At R_x the value of $\mathbf{H}'$ is $\mathbf{H}_c'$, the coercive force of a ring. Thus

$$R_x=0.2\frac{NI}{-\mathbf{H}_c'} \qquad (4)$$

Also, by definition

$$\mathbf{H}=\frac{0.2NI}{1/2(R_2+R_1)} \qquad (5)$$

The solution for **B** as a function of **H** is now readily obtained

$$\begin{aligned}\mathbf{B}&=\frac{1}{R_2-R_1}[-\mathbf{B}_m(R_x-R_1)+\mathbf{B}_m(R_2-R_x)]\\&=\frac{1}{R_2-R_1}[\mathbf{B}_m(R_2+R_1)+2\mathbf{B}_m\frac{\mathbf{H}}{\mathbf{H}_c'}\times\\&\qquad 1/2(R_2+R_1)]\end{aligned}$$

$$\mathbf{B}=\frac{R_2+R_1}{R_2-R_1}\left(1+\frac{\mathbf{H}}{\mathbf{H}_c'}\right)\mathbf{B}_m \qquad (6)$$

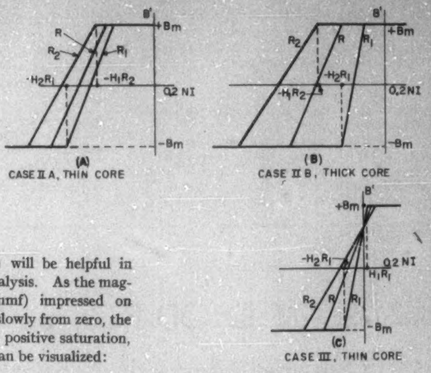


Fig. 2. Flux versus mmf for elementary rings

A—Case II-A, thin core
B—Case II-B, thick core
C—Case III, thin core

The slope of the side of the loop is then

$$\frac{d\mathbf{B}}{d\mathbf{H}}=\mu_d=\frac{R_2+R_1}{R_2-R_1}\frac{\mathbf{B}_m}{\mathbf{H}_c'} \qquad (7)$$

μ_d as defined in the foregoing is termed the differential permeability of the core.

The meanings of equations 6 and 7 are simple enough. The coercive force for the core is always $\mathbf{H}_c'$, independent of the ID/OD ratio. On the other hand, the flank is no longer vertical but is a straight line having the slope

$$\mu_d=\frac{1+r}{1-r}\frac{\mathbf{B}_m}{\mathbf{H}_c'} \qquad (8)$$

where

$$r=\frac{R_1}{R_2}=\text{ID/OD}$$

From equation 8 it is seen that reducing r from 1 (infinitely thin core) first results in a rapid decrease in μ_d, and then in a gradual approach to the minimum value $\mathbf{B}_m/\mathbf{H}_c'$ (infinitely thick core). It will be shown later that experimental results on cores of oriented 50-per-cent nickel iron follow this relation reasonably accurately.

CASE II. PARALLELOGRAM LOOP WITH $\mathbf{B}_r=\mathbf{B}_m$

The problem to be solved here is the same as that stated under case I. Equations 1, 3, and 5 are still applicable. The $\mathbf{B}'$-$\mathbf{H}'$ relation for the thin rings is as shown in Fig. 1(B)

$$\begin{aligned}\mathbf{B}'&=+\mathbf{B}_m && -\mathbf{H}_1<\mathbf{H}'\\&=\mu_d'(\mathbf{H}'+\mathbf{H}_c'), && -\mathbf{H}_2<\mathbf{H}'<-\mathbf{H}_1\\&=-\mathbf{B}_m && \mathbf{H}'<-\mathbf{H}_2\end{aligned} \qquad (9)$$

Consideration of the physical picture discloses the existence of two subcases, (A) thin core, and (B) thick core. In case II-A, a range of mmf values exists for which none of the rings of the core are saturated. Fig. 2(A) shows $\mathbf{B}'-NI$ curves for the rings at the inner, outer, and an intermediate radius; from it, the unsaturated range is seen to be

$$-\mathbf{H}_1R_2>0.2NI>-\mathbf{H}_2R_1$$

It will be shown in the following that the **B-H** relation is linear over this range, which may be large or small depending upon both the geometry of the core and the core material properties. The remainder of the left flank consists of upper and lower nonlinear regions fairing smoothly from the linear to the saturated regions.

In case II-B, thick core, no unsaturated range exists, as can be seen from Fig. 2(B). Nevertheless, a linear region in the **B-H** relation is found in the range

$$-\mathbf{H}_2R_1>0.2NI>-\mathbf{H}_1R_2$$

Nonlinear regions the same as those in case II-A complete the curve.

The criterion determining under which of these two cases a core falls is found from the foregoing inequalities. Thus, for case II-A

$$-\mathbf{H}_1R_2>-\mathbf{H}_2R_1$$

$$\frac{\mathbf{H}_1}{\mathbf{H}_2}<\frac{R_1}{R_2}$$

Introducing the symbols $h=\mathbf{H}_1/\mathbf{H}_2$ and $r=R_1/R_2$

$$h<r \tag{10}$$

Similarly for case II-B

$$h>r \tag{11}$$

The boundary case $h=r$ is of mathematical interest only. The linear region of the **B-H** relation has zero length, and the entire curve is made up of the two nonlinear regions. A graph of the equations for this case for a practical value of h was made to determine the degree of nonlinearity. The result was that a straight line fit could be made to approximately 80 per cent of the left flank, i.e., the theoretically nonlinear functions have very good practical linearity. No further treatment of the nonlinear regions will be given in this paper.

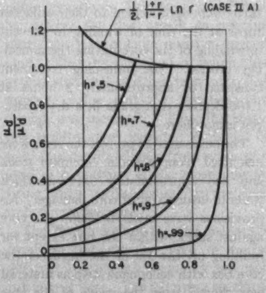


Fig. 3. Effects of ID/OD ratio and material characteristics on differential permeability, cases II-A and II-B

Case II-A. Thin Core, Linear Region

From equation 1

$$\mathbf{B}=\frac{1}{R_2-R_1}\int_{R_1}^{R_2}\mu_d'(\mathbf{H}'+\mathbf{H}_c')dR$$

$$=\frac{\mu_d'}{R_2-R_1}\left[0.2NI\ Ln\frac{R_2}{R_1}+\mathbf{H}_c'(R_2-R_1)\right]$$

$$\mathbf{B}=1/2\mu_d'\left(\frac{1+r}{1-r}\right)\left(Ln\frac{1}{r}\right)\mathbf{H}+\mu_d'\mathbf{H}_c' \tag{12}$$

The ratio of the slope of this region μ_d to the slope of the flank of the thin core loop μ_d' is then

$$\frac{\mu_d}{\mu_d'}=\frac{1}{\mu_d'}\frac{d\mathbf{B}}{d\mathbf{H}}=1/2\left(\frac{1+r}{1-r}\right)Ln\frac{1}{r} \tag{13}$$

This function is plotted in Fig. 3. It is seen to be essentially constant for a wide range of r and is independent of h. The unexpected increase of μ_d over μ_d' is because of the arbitrary definition of **H** as the value of field intensity at the mean radius rather than at some other radius.

It can easily be shown that the coercive force $\mathbf{H}_c$ for the core is given by

$$\frac{\mathbf{H}_c}{\mathbf{H}_c'}=\left(\frac{\mu_d}{\mu_d'}\right)^{-1} \tag{14}$$

Thus, while the coercive force is no longer perfectly constant as it was in case I, it is essentially constant in the range of r covered by the experiments reported here.

Case II-B. Thick Core, Linear Region

Taking into account the rings in the core which for a given NI are positively saturated, unsaturated, or negatively saturated, equation 1 yields

$$\mathbf{B}=\frac{1}{R_2-R_1}\left[\mathbf{B}_m\left(R_2+\frac{0.2NI}{\mathbf{H}_1}\right)-\mathbf{B}_m\left(\frac{-0.2NI}{\mathbf{H}_2}-R_1\right)+\int_{\frac{0.2NI}{\mathbf{H}_2}}^{\frac{0.2NI}{\mathbf{H}_1}}\mu_d'\left(\mathbf{H}_c'+\frac{0.2NI}{R}\right)dR\right]$$

Integrating and simplifying, the result is

$$\mathbf{B}=1/2\mu_d'\left(\frac{1+r}{1-r}\right)\left[(\mathbf{H}_2-\mathbf{H}_1)+\mathbf{H}\ Ln\frac{1}{h}\right] \tag{15}$$

$$\frac{\mu_d}{\mu_d'}=1/2\left(\frac{1+r}{1-r}\right)Ln\frac{1}{h} \tag{16}$$

The function of equation 16 is plotted in Fig. 3 for several values of h. Comparison of equation 16 with 13 shows that the only difference is the replacement of $Ln(1/r)$ by $Ln(1/h)$. Thus, the two curves join at a break point which represents the special case $r=h$. Note that Fig. 3 applies only to the linear regions of the **B-H** relation and that in the vicinity of the break point these regions are theoretically small. The decrease of μ_d with decreasing r is steep at first and then more gradual as the minimum value $1/2\ \mu_d'\ Ln(1/h)$ is approached ($r=0$).

In terms of the magnetic properties of the elementary rings, the break point r_b is given by

$$r_b=h=\frac{\mathbf{H}_c'-\mathbf{B}_m/\mu_d'}{\mathbf{H}_c'+\mathbf{B}_m/\mu_d'}=\frac{\mu_d'\mathbf{H}_c'-\mathbf{B}_m}{\mu_d'\mathbf{H}_c'+\mathbf{B}_m} \tag{17}$$

The coercive force is constant as r varies and is

$$\mathbf{H}_c=\frac{\mathbf{H}_1-\mathbf{H}_2}{Ln\frac{1}{h}}=\left[1/2\mu_d'\left(\frac{1+h}{1-h}\right)Ln\frac{1}{h}\right]^{-1} \tag{18}$$

This value is equal to that given by equation 14 at the break point, as it must be.

Case III. Parallelogram Loop with $\mathbf{B}_r<\mathbf{B}_m$

This case, shown in Fig. 2(C), is different from case II in that there is no thick core subcase. The equation for μ_d in the linear region defined by

$$-\mathbf{H}_2R_1<0.2NI<\mathbf{H}_1R_1$$

is exactly the same as for case II-A, equation 13. Therefore, μ_d is always greater than μ_d'. The coercive force is given by equation 14.

Discussion

Practical core materials do not have perfect parallelogram hysteresis loops. A brief investigation was made of the effects of rounded corners by means of calculations based on a polyhedral loop. The apparent major result was a reduced value of effective $\mathbf{B}_m$, which indicates that, in any quantitative calculation of the effect of the ID/OD ratio on practical core materials, it might be desirable to use a value of $\mathbf{B}_m$ less than the measured value for the material.

Experimental

Experimental Approach

The technique used to investigate experimentally the influence of core ID/OD ratio on static and effective dynamic core magnetic properties was as follows: Using a nickel-iron alloy extensively employed in magnetic amplifiers, a group of cores was manufactured having as large a radial buildup (small ID/OD ratio) as possible. These cores were placed in specially constructed high-rigidity protective boxes; the necessary windings were applied; and electrical tests were made to determine the desired magnetic properties. Then the windings were cut off, the boxes opened and, with extreme caution, a portion of the core was removed by clipping wraps of strip from the inside of the core. The resultant cores, having an increased ID/OD ratio but unchanged basic material properties, were then retested for net magnetic properties. This procedure was repeated several times until there was so little core material left that electrical testing became impractical.

The set of electrical measurements made at each step in the reduction of core size consisted of the following: flux-meter measurements of the left flank of the major static hysteresis loop, control characteristic curve of half-wave self-saturating amplifier with a 60-cycle-per-second (cps) supply frequency, control characteristic curve of half-wave self-saturating amplifier with 400-cps supply frequency, and constant-current flux-reset core test[5] characteristic curve for a 400-cps test frequency.

Data from the half-wave amplifier and core tester measurements were used to calculate the important features of corresponding control magnetization curves. Using the analogy between the control magnetization curve and an effective major dynamic hysteresis loop, values of effective magnetic quantities such as coercive force, saturation flux density, residual flux density, and differential permeability can be defined. The effective magnetic quantities calculated from the half-wave amplifier data were coercive force and differential permeability. From core tester measurements, values of saturation and residual flux density were found in addition to coercive force and differential permeability. By plotting the experimentally determined values of static and effective dynamic magnetic quantities as a function of the ID/OD ratio, a comparison was made with the theoretically derived relationships.

It was expected that if the approach taken were valid then the plots of magnetic quantities versus the ID/OD ratio for the various types of electrical measurements would compare with the theoretical expressions as follows: Static hysteresis loops, for the type of magnetic alloy being considered, have nearly vertical sides; thus the static magnetic quantities should vary with the ID/OD ratio according to the relations derived under the theoretical case I (perfectly rectangular loop). Control magnetization curves for half-wave amplifier operation or for core tester operation have finite slopes or differential permeabilities even for cores with very small radial buildup. Further, the values of differential permeability are known to decrease with increasing operating frequency. Thus, effective dynamic magnetic quantities should vary with the ID/OD ratio approximately according to the relations derived under the theoretical case II (parallelogram loop with $\mathbf{B}_r = \mathbf{B}_m$), and the results for the two different test frequencies should be similar except for a shift in the basic material values of $\mathbf{H}_c'$ and μ_d'. The results from the 400-cps core tester measurements were expected to be nearly identical to those from the 400-cps half-wave amplifier measurements. They were included both as a further check on the correlation between core tester and amplifier operation, and as a check on the applicability of the theoretical results under slightly altered core operating conditions.

It is apparent that with the procedure described in the foregoing for experimentally varying the ID/OD ratio, there are several possible sources of error or variance of core properties not attributable to ID/OD variation. First, it will be noted that each cut in core size changes both the core cross-sectional area and mean magnetic path length as well as the ID/OD ratio. So long as the flux paths remain entirely within the core material, the effect of variations may be easily taken into account in the conversion of electrical data to magnetic quantities. This is a valid assumption for the present case of high-permeability core materials and strip-wound cores having a large number of wraps. In general, it is found that for more than about 10 wraps in a core, the end effects are completely negligible and the flux paths are for all practical purposes in the plane of the strip.

Some other possible sources of error inherent in the technique used are:

1. Mechanical damage to the remaining portions of the core material when wraps are clipped from the inside of the core.
2. Variations in the basic magnetic material in making the core.
3. The presence of nonuniformly distributed localized strains resulting perhaps from interlaminar welds caused by poor strip-insulating coating.
4. Metering errors in obtaining the electrical data.

The influence of errors such as these on the validity of the comparisons to be made in the course of the experiment and the conclusions to be drawn therefrom can be reduced by carrying a group of nearly identical cores through the experiment concurrently and also by taking as large a body of electrical data as practical. The presence of errors can then be detected by a divergence of results.

Core Details

The magnetic alloy chosen for the experimental investigation was Hipernik V, a grain-oriented 50-per-cent nickel-iron alloy similar in properties to Deltamax, Orthonol, Orthonik, etc. This material was used because of its ready availability and widespread use in magnetic amplifiers. The particular strip selected was of 2-mil thickness and 1/4-inch width and was chosen from a number of available batches or melts as having the desirable characteristics of being relatively insensitive to minor variations in annealing treatment and of normally yielding cores having a small variance of magnetic properties. The static magnetic properties normally obtained with this material are $\mathbf{H}_c = 0.11$ oersted, $\mathbf{B}_m = 14.6$ kilogausses, and $\mathbf{B}_r/\mathbf{B}_m = 0.96$ for a tip $\mathbf{H}$ of 2 oersteds.

Four cores were wound from this strip to an over-all size of 1.90 centimeters ID by 5.5 centimeters OD by 0.635 centimeter height with approximately an 80-per-cent space factor. An insulating material consisting of MgO plus adhesion-promoting additives was applied to one side of the strip before winding into cores. The approximate weight of each core was 90 grams. Both the inner and outer turns were spot-welded to their adjacent turns at the time of winding to prevent unwinding of the core during the annealing process. The annealing treatment consisted of approximately 2 hours at 1,100 degrees centigrade in a dry hydrogen atmosphere.

The boxes used to protect the cores were machined from phenolic laminate stock and had approximately 1/16-inch-thick walls to insure against core damage. No provisions were made to damp the core against shock within the box except for building up the ID of the inside of the core box with cellophane strip as material was clipped from the inside of the core during the course of the experiment. By thus building up the core box to follow the changes in core dimension, it was pos-

Vo
Pt.
19

Table I. Steps in Reduction of Core Size

Cut No.	r_{av}
0	0.35
1	0.43
2	0.51
3	0.58
4	0.65
5	0.72
6	0.81
7	0.91
8	0.97

sible to use the same boxes throughout the experiment except for the final cut in core size which required a new box with a larger center hole to accommodate the test windings.

In all, apparent magnetic properties were obtained for each core at nine different values of the ID/OD ratio. An approximately equal amount of material was removed from each core at each cutting. However, there were variations in the amount of material removed per cut since this was gauged by eye and was of secondary importance with respect to taking care to minimize damage to the remaining core material. Listing the steps in reduction of core size by cut number, running from 0 to 8, the average ID/OD ratio r_{av} for the group of four cores corresponding to each cut number is listed in Table I.

ELECTRICAL MEASUREMENTS

Static **B-H** loop measurements were made with a standard commercially available fluxmeter (not a ballistic galvanometer) but in a somewhat unconventional manner. To maintain the same mode of core operation as in a magnetic amplifier down to zero frequency, the cores were operated statically over a series of minor hysteresis loops similar in form to those traversed in magnetic amplifier operation. This was accomplished by driving the flux nearly to positive saturation with a large positive applied mmf. Then by means of a switch, the applied mmf was switched to a controlled negative value and the resulting flux change was measured. Thus the static test results were presented in terms of total flux density change from saturation level as a function of the mean magnetizing force corresponding to the controlled negative mmf.

The design values for the half-wave amplifier test circuits were of course different for each test frequency and each cut in core size. The following specifications were used in the design of these amplifier circuits:

1. A-c supply voltage for the amplifier to be sinusoidal with less than 10-per-cent total harmonic distortion and magnitude within ±5 per cent of the saturation voltage of the gate winding. The saturation voltage is here defined as that voltage which with the rectifier short-circuited will cause a magnetizing current to flow whose instantaneous peak value is twice the value at the 90-degree point, i.e., at the peak of applied voltage.
2. Self-saturating rectifier to be a vacuum tube so selected in conjunction with the circuit design that the peak rectifier forward drop under any condition of amplifier control is less than 20 per cent of the a-c supply voltage.
3. The amplifier load to be resistive and of a value such that the maximum value of d-c load current is within 20 per cent of 20 times the minimum value of d-c load current obtained at amplifier cutoff.
4. The control current to be well filtered d-c obtained from a high-impedance source such that the peak deviation from the average, caused either by ripple from the source or voltage induced in the control winding of the amplifier, does not exceed 0.25 per cent under any conditions over the normal control region of the amplifier.

The half-wave amplifier control characteristic curves were taken with the aid of an automatic 2-axis plotter in the form of per-cent load voltage or current versus control ampere-turns. Maximum or 100-per-cent output was obtained by placing a large positive control mmf on the core. The value of a-c supply voltage was also recorded. Calculation of the control magnetization curve from this amplifier data is covered in the next section.

Constant-current flux-reset core tester measurements were made as described in reference 5. Briefly, an excitation mmf having a half-wave rectified wave form and a peak value sufficient to drive the flux to saturation is applied to the core. A separate d-c control mmf is also applied which, when negative, resets the flux during the zero half-cycles of excitation mmf. A special average-sensitive meter is used to measure the induced voltage in a separate pickup winding on the core. The test consists of measuring the induced voltage as a function of the control mmf. Details of the conversion of this type of data to a control magnetization curve are given also in reference 5.

CONTROL MAGNETIZATION CURVE CALCULATIONS

Following the methods given in reference 5, it can be shown that, to a good approximation, the direct voltage V_L appearing across the total gate circuit resistance of a half-wave amplifier with resistive load and rms supply voltage V_{AC} will be given by

$$V_L=\frac{\sqrt{2}}{\pi}\left(V_{AC}-V_G\right) \tag{19}$$

where V_G is the a-c induced voltage in the gate winding as measured on an average-sensitive a-c meter calibrated in rms for a sine wave. V_G will be related to the peak-to-peak cyclic flux density swing in gausses by

$$V_G=\frac{\pi}{\sqrt{2}}\Delta\mathbf{B}fA_CN_G10^{-8} \tag{20}$$

where

f = frequency

A_C = effective cross sectional area, square centimeters

N_G = gate winding turns

Solving equation 19 for V_G

$$V_G=V_{AC}\left(1-\frac{\pi}{\sqrt{2}}\frac{V_L}{V_{AC}}\right) \tag{21}$$

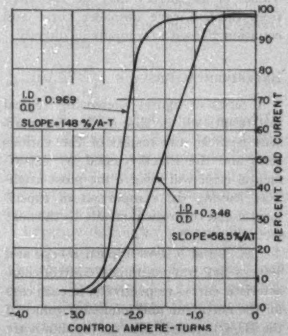


Fig. 4. Typical 60-cps half-wave amplifier characteristics for minimum and maximum ID/OD ratio

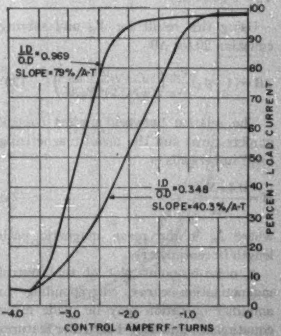


Fig. 5. Typical 400-cps half-wave amplifier characteristics for minimum and maximum ID/OD ratio

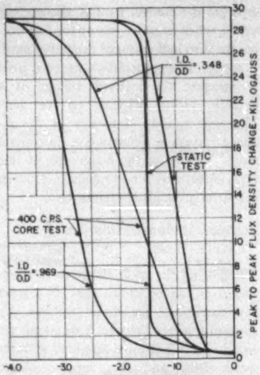


Fig. 6. Typical static and 400-cps core test data for minimum and maximum ID/OD ratio

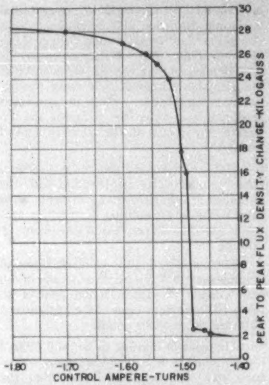


Fig. 7. Expanded portion of d-c hysteresis loop shown in Fig. 6

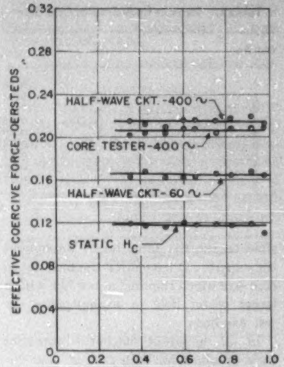


Fig. 8. Typical coercive force versus r plots

However, it is not convenient to measure V_L because this is the voltage across the total circuit resistance including the winding and rectifier forward resistances. It is possible and more convenient to express V_G as a function of the per-unit output $\mathscr{g}_L$ rather than V_L. Noting that the maximum value of V_L obtained when the core is saturated at all times is $V_{AC}\sqrt{2}/\pi$, it will be seen that

$$\mathscr{g}_L=\frac{V_L}{V_{L(\max)}}\ \frac{V_L}{V_{AC}\sqrt{2}/\pi} \qquad (22)$$

$\mathscr{g}_L$ is defined by equation 22 in terms of load voltage but could obviously refer to per-unit output current as well.

Substituting equation 22 into equation 21

$$V_G=V_{AC}(1-\mathscr{g}_L)$$

Using this result for V_G and solving equation 20 for $\Delta\mathbf{B}$

$$\Delta\mathbf{B}=(1-\mathscr{g}_L)\frac{V_{AC}}{N_G}\left(\frac{10^8}{(\pi/\sqrt{2})fN_GA_C}\right) \qquad (23)$$

The relation between applied control ampere-turns and the mean magnetizing force in oersteds is

$$\mathbf{H}=\frac{0.4\pi N_cI_c}{l_m} \qquad (24)$$

where l_m is the mean magnetic path length in centimeters.

A point-by-point plot of the control magnetization curve corresponding to amplifier operation may be made using equations 23 and 24. Particular features of the control magnetization curve which are of interest include the effective values of coercive force and differential permeability in the linear region of the curve. The values of effective coercive force were determined directly from the half-wave amplifier control characteristic curve by using equation 23 in reverse to determine the value of $\mathscr{g}_L$ for $\Delta\mathbf{B}=\mathbf{B}_m$, and converting this to oersteds by equation 24. The effective differential permeability at any point on the control magnetization curve is defined by

$$\mu_d=\frac{d(\Delta\mathbf{B})}{d\mathbf{H}}$$

Substituting in values of $\Delta\mathbf{B}$ and $\mathbf{H}$ from equations 23 and 24

$$\mu_d=\frac{d\mathscr{g}_L}{d(NI)}\ \frac{V_{AC}}{N_G}\ \frac{10^8\,l_m}{2\sqrt{2}\pi^2fA_C} \qquad (25)$$

Values of $d\mathscr{g}_L/d(NI)$ may be measured from the half-wave amplifier curve and equation 25 used to compute μ_d directly.

EXPERIMENTAL RESULTS

A series of figures presenting typical test results will be discussed. Comparisons between the results of the various tests and the effects caused by experimental error will receive the most attention, leaving the comparison of experimental and theoretical results to the next section.

Figs. 4 and 5 show typical 60-cps and 400-cps half-wave amplifier control characteristic curves respectively, in each case for the maximum and minimum values of the ID/OD ratio. The curves shown are for one of the four cores tested, but results for the others were very similar. The increase of slope because of increased ID/OD ratio is evident at both frequencies but is greater at 60 cps because of the higher value of μ_d' which results from smaller eddy-current effects at the lower frequency. The value of control current for cutoff, i.e., the ampere-turns necessary to reset completely the outside ring of the core, remains approximately independent of the ID/OD ratio because the latter was varied by removing wraps from the inside of the core. The nonlinearity apparent at low values of output is caused by backfiring[8] within the gate circuit of the amplifier during reset. While the reset wave form was observed to change somewhat as the ID/OD ratio was varied, backfiring is primarily caused by the characteristics of the material used. The backfiring effect completely masked any linearity variations caused by varying ID/OD ratio.

The typical 400-cps core test curves of Fig. 6 are seen to be closely similar to the corresponding amplifier curves of Fig. 5. The static test curves, also in Fig. 6, exhibit a higher slope than the dynamic test curves. For very thin cores, the slope was found to be extremely high; an enlarged portion of such a hysteresis loop is shown in Fig. 7. Accurate d-c measurements for thin cores were impossible to make because of the extreme sensitivity of the core to mechanical vibration such as a light tapping of the core box during the measurement. The data points of Fig. 7 were therefore found by averaging a number of trial readings. Aside from the disturbances caused by vibration, the flank of the loop also exhibits definite steps, two of which can be discerned in Fig. 7. Observation of the fluxmeter during a change of flux

Table II. Summary and Comparison of Experimental Data

	Average Measured Values			Values Determined From μ_d Versus r Data			Calculated Values		
	H_c, Oersteds	B_m, Kilogausses	B_m-B_r, Kilogausses	μ_d'	K	r_b	K	r_b	B_m, Kilogausses*
Static core tests									
Core *A*	0.118	14.81	0.46		118,000	1.000	125,500	1.000	13.80
Core *B*	0.114	14.93	0.54		126,000	1.000	131,000	1.000	14.38
Core *C*	0.112	14.86	0.49		137,000	1.000	132,500	1.000	15.30
Core *D*	0.122	14.85	0.47		132,000	1.000	122,000	1.000	16.10
Average	0.116	14.86	0.49		128,000	1.000	127,700	1.000	14.92
60-cps amplifier tests									
Core *A*	0.163	†	‡	632,000	76,500	0.790	90,000	0.702	12.05
Core *B*	0.163	†	‡	577,000	72,000	0.772	90,400	0.732	12.10
Core *C*	0.157	†	‡	700,000	79,500	0.795	90,350	0.765	12.50
Core *D*	0.176	†	‡	580,000	74,200	0.770	83,200	0.750	13.20
Average	0.165			622,000	75,600	0.782	88,500	0.737	12.46
400-cps core tests									
Core *A*	0.206	14.61	0.48	420,000	58,200	0.759	71,700	0.711	11.85
Core *B*	0.209	14.56	0.513	375,000	60,500	0.722	70,400	0.686	12.65
Core *C*	0.198	14.70	0.48	452,000	68,700	0.770	74,800	0.717	11.60
Core *D*	0.223	14.68	0.44	472,000	59,500	0.770	66,200	0.755	13.60
Average	0.209	14.64	0.48	430,000	61,700	0.755	70,800	0.717	12.43
400-cps amplifier tests									
Core *A*	0.213	†	‡	395,000	51,900	0.770	69,000	0.705	11.00
Core *B*	0.214	†	‡	375,000	56,000	0.722	68,800	0.693	12.95
Core *C*	0.204	†	‡	400,000	60,200	0.762	72,300	0.693	11.00
Core *D*	0.228	†	‡	378,000	56,400	0.745	64,500	0.710	12.60
Average	0.215			387,000	56,100	0.750	68,600	0.700	11.89

*Corrected value of B_m necessary for equality of calculated and measured values of *K*.
†Saturation flux density not measured under amplifier operation; for calculations, $B_m = 14.6$ kilogausses assumed.
‡Not accurately determined from amplifier measurements because of test equipment deficiencies.

showed that the change was not smooth but consisted of a series of small jumps. The total time required to reach a steady-state flux level was in some cases as long as several minutes. A possible explanation of the steps observed is that the main boundary walls within the core material are impeded or "hung up" on isolated impurities within the crystal lattice.

Effective values of coercive force for a typical core as a function of the ID/OD ratio are plotted in Fig. 8 for all of the test methods. In each case the coercive force is essentially constant and, as expected, it increases with test frequency. The small difference in coercive force between the amplifier and the core test measurements at 400 cps is probably because of backfiring in the amplifier. Coercive force curves for the other three cores were essentially the same as for core *A*. Average values for all of the cores can be found in Table II.

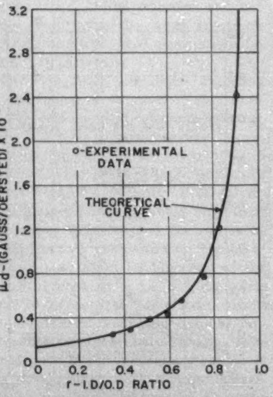


Fig. 9. Typical μ_d versus r plot for static test, core A

Differential permeability versus ID/OD ratio for typical static test data are shown in Fig. 9. The individual data points shown correspond to actual measurements of the average slope of the static loops, while the solid curve is a plot of the theoretical relation

$$\mu_d = K\left(\frac{1+r}{1-r}\right) \qquad (26)$$

where the value of *K* is determined for best fit of the data. Similar plots for the other cores were very nearly identical except for slight variations in the value of *K* for best fit. Values of *K* for all cores can be found in Table II.

Effective differential permeability versus ID/OD ratio for the dynamic tests is plotted for each of the four cores tested in Figs. 10 through 13. Here as before the individual points represent actual measurements while the solid curves are plots of the theoretical relation with the proportionality constant *K* chosen for best fit. It will be noted that closeness of fit varies somewhat from curve to curve. The accuracy of the data is, however, uniformly poorer for high values of ID/OD ratio because of several factors. First, higher values of amplifier or core tester gain become increasingly difficult to measure accurately, particularly when the data are taken with an automatic curve plotter. Second, thinner cores become more and more delicate and therefore sensitive to mechanical stress; however, other test data such as core tester measurements of $B_m - B_r$ give little evidence of straining.

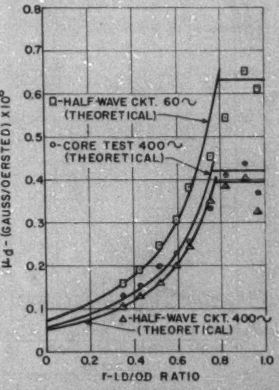


Fig. 10. μ_d versus r for dynamic tests, core A

Discussion and Conclusions

The importance of investigating the effect of the ID/OD ratio on core gain was first realized when attempts were made to calculate optimum core dimensions for magnetic amplifier applications. All the other factors affecting amplifier performance such as winding space factor and winding resistance could be approximated satisfactorily. It was known from tests on two cores of widely different ID/OD ratios only that the effect on gain was by no means negligible, being more than 50 per cent in this instance.

The theoretical results presented graphically in Fig. 3 were then derived on the basis of the simplest reasonable assumptions, as described earlier in this paper. The physical existence of the plateau and the sharp break predicted by the curves were difficult to accept, probably because of a fundamental belief in the continuity of nature. Further, the desire to prove that an error had been made was fostered by the realization that optimization calculations would not be simple if such a nonlinear relation had to be used.

It was therefore felt necessary to set up the rather elaborate experiment described in the foregoing to get reliable data and to prove or disprove the existence of the break. Examination of the results, Figs. 10 through 13, shows that the influence of the ID/OD ratio is indeed of the form which was predicted. In spite of the unavoidable dispersion of values measured near $r=1$, it is reasonably clear that the data confirm rather than deny the existence of a break point and plateau. The static test results, Fig. 9, verify the theoretical equation 8, which was also given by McConnell.[1]

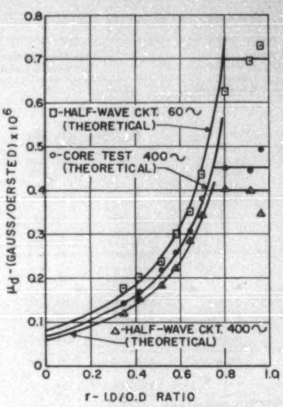


Fig. 12. μ_d versus r for dynamic tests, core C

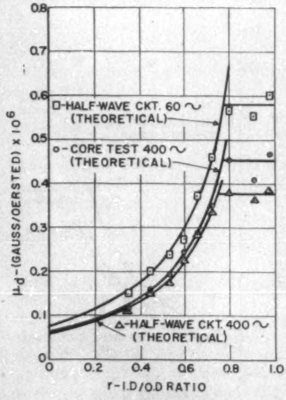


Fig. 11. μ_d versus r for dynamic tests, core B

A more detailed quantitative comparison of theory and experiment can be made with the aid of Table II, which lists measured values of $\mathbf{H}_c$, $\mathbf{B}_m$, μ_d', K, and R_b for each type of test on each core. The values of μ_d', K, and r_b from the μ_d versus r plots were for the best fit of the theoretical equation to the data points. That is, K was chosen for best fit in the thick core region, μ_d' was the average of data points in the thin core (plateau) region, and r_b was read from the intersection of the plateau and the curve. Table II also lists calculated values of r_b, equation 17, and

$$K=1/2\mu_d' \, Ln\left(\frac{1}{h}\right) \tag{27}$$

from equations 16 and 26. The individual average values for $\mathbf{H}_c'$ and $\mathbf{B}_m$ and the measured value for μ_d' (the only available value) were used in the calculations. Agreement of measured and calculated values of K and r_b is quite good for the static tests, but the calculated K values are approximately 10 per cent higher for the dynamic tests. It was thought that this discrepancy might be caused by the rounding of the top and bottom of the a-c control magnetization curves for thin rings. As mentioned in the section entitled "Derivation of Equations," the effect of this rounding was found to be a reduction in the effective value of $\mathbf{B}_m$. Following this suggestion, values of $\mathbf{B}_m$ which lead to equality of measured and calculated values of K were found and are listed in Table II. It appears that on the average a value of approximately 12 kilogausses results for $\mathbf{B}_m$. This value is physically reasonable since it corresponds roughly to the knee of the a-c control magnetization curves.

Although this work has been restricted to toroidal cores of a single type of magnetic alloy, a certain amount of information is available to indicate that similar results can be expected for other core shapes and magnetic materials.

An indication of the presence of an appreciable ID/OD ratio effect on the magnetic amplifier performance of stacked DU laminations of silicon-iron has been given by Teasdale and Brownell.[6] They found that reducing the width of the legs of the U-shaped laminations increased core gain as measured on a constant-current flux-reset core tester.

A direct consequence of the theoretical results shown in Fig. 3 is that, to realize maximum gain from a material having a high value of h, it is necessary to make $r>h$. The resulting core shape may have smaller radial build-up than is ordinarily found in commercially available cores. Limited experimentation with cores of Hymu-80 alloy, a Mo-Permalloy material, disclosed that, on the basis of core tester measurements at 400 cycles, a change in the ID/OD ratio from 0.66 to 0.80 caused effective μ_d to change from 500,000 to 1,000,000 while a further increase of r to 0.89 gave a negligible change in μ_d. These results were obtained with separate groups of cores wound to the various values of the ID/OD ratio, rather than the technique used to vary ID/OD previously described. It is of interest to note that immediate practical use was

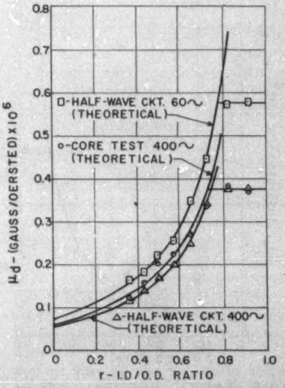


Fig. 13. μ_d versus r for dynamic tests, core D

made of the increased μ_d available with the increased ID/OD ratio. It was possible to double the inherent gain of the first stage of a 2-stage all-magnetic servo amplifier by the simple expedient of doubling the strip width and halving the radial build-up of the first stage cores.

When considering the application of the results found here to the design of cores for magnetic amplifiers, the first thought is to restrict core radial build-up so as to be always in the thin core region. However, this is not necessarily an optimum condition for all applications. Other considerations such as output power and toroidal winding restrictions may well outweigh the influence of the ID/OD ratio on basic core gain. However, in the design of a core for any particular application, the influence of core ID/OD ratio must certainly be known and taken into account.

References

1. Static Properties of Toroidal Magnetic Cores, H. M. McConnell. *Magnetic Amplifier Technical Report No. 13*, Carnegie Institute of Technology, Pittsburgh, Pa., 1953.
2. Evaluation of Core Materials for Magnetic Amplifiers, Part I, D. C. Dieterly. *Electrical Manufacturing*, New York, N. Y., Jan. 1953.
3. Predetermination of Control Characteristics of Half-Wave Self-Saturating Magnetic Amplifiers, H. Lehmann. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2097–103.
4. Evaluation of Core Materials for Magnetic Amplifiers, Part II, D. C. Dieterly. *Electrical Manufacturing*, New York, N. Y., Feb. 1953.
5. Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers, R. W. Roberts. *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1953 section), pp. 682–90.
6. Evaluation of Core Materials for Magnetic Amplifier Applications, R. D. Teasdale, H. R. Brownell. 1954 National Convention Record, pt. III, Mar. 1954, pp. 56–65.

Oscillographic Techniques for the Evaluation of Magnetic Amplifier Response

D. L. CRITCHLOW
STUDENT MEMBER AIEE

A COMMON way to study the transient response of a magnetic amplifier upon application of a step of signal or to examine its frequency response to periodical signals is to use a recording oscillograph or a cathode-ray oscilloscope. If an oscilloscope is used its sweep rate can be slowed down so as to contain a whole transient in a single sweep. Repetitive switching arrangements controlling the initiation of the transient and the time sweep make the latter procedure suitable for direct visual observation of a stationary pattern.[1]

However, these procedures, appealing as they may be in view of their simplicity, seldom yield adequate information. Changes of output voltage or current are judged from the observation of an envelope of peaks of half-cyclic wave forms while the quantity of actual interest is not the peak value but rather the half-cyclic average or the rms value of the output. In particular, the envelope of peaks seen for any part of the transient when the angle of firing is less than 90 degrees appears to have constant amplitude even though the output quantity of interest is actually varying. The recognition of this inadequacy has led to the development of "half-cyclic integrators" which supply information about the half-cyclic average of the output throughout a transient[2] or in the steady state with periodic signals.[3] However, such integrators involve the construction of special electronic gear of rather high accuracy. This limits their acceptance in practice. In this paper some simple techniques are described which have been used successfully at the Carnegie Institute of Technology by employing oscilloscopes and other components of standard design commonly found in the laboratory.

Transient Response

The half-cyclic output of magnetic amplifiers with sensitive core materials is readily expressed in terms of the angle of firing. During a transient this angle varies from half-cycle to half-cycle, proceeding from the initial to the final steady-state value. Superposed records of half-cycles of output throughout the transient are obtained easily with an oscilloscope by synchronizing the time sweep externally with the voltage supply frequency; proper sweep rates and horizontal gains allow a display of a pattern of subsequent half-cycles or of half-cycles equally spaced in time (e.g., one half-cycle each *k* cycles). The pattern can be recorded photographically or can be observed directly if a long-persistence screen or a repetitive switching arrangement of the kind described in reference 1 is used.

Resistive Loads

Fig. 1(A) shows one of these patterns of output currents for a center-tap amplifier with resistive load. The gate voltage frequency is 60 cycles. Since the tran-

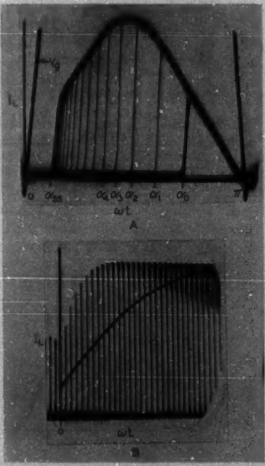


Fig. 1. Transient response of center-tap amplifier with resistive load

A—Superposed traces, k = 2
B—Envelope of peaks of output current; full line in the response calculated from A

Paper 55-545, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 28, 1955; made available for printing May 5, 1955.

D. L. Critchlow is with the Carnegie Institute of Technology, Pittsburgh, Pa.

This work was supported by the U. S. Office of Naval Research. Suggestions and comments by L. A. Finzi, Project Supervisor, are acknowledged.

Table I. Calculation of Transient Response: Resistive Load

Trace No., n	Cycle No., kn	Firing Angle α, Degrees	$1+\cos\alpha$	I_{av}, Milliamperes
0	0	129	0.37	17.4
1	2	107	0.71	33.2
2	4	88	1.04	48.5
3	6	74	1.28	59.7
4	8	64	1.44	67.1
5	10	55	1.57	73.2
6	12	48	1.67	77.8
8	16	38	1.79	83.5
10	20	31	1.86	86.6
∞		21	1.93	90.0

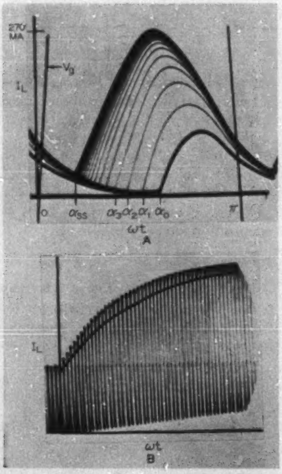


Fig. 2. Transient response of center-tap amplifier with resistive inductive rectified load

A—Superposed traces, k=3
B—Envelope of peaks of output current; full line is the response calculated from A

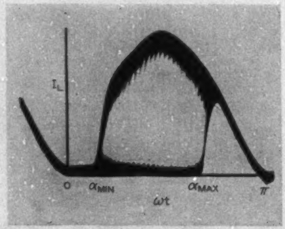


Fig. 3. Response of a 400-cycle center-tap amplifier with d-c bias and a 30-cycle control signal. Measurement of minimum and maximum outputs

sient in question was known to last about 25 cycles, only one half-cycle every 2 cycles has been displayed, i.e., $k=2$. In this manner a confusing overcrowding of traces has been avoided and yet enough information has been obtained concerning the changes of output throughout the transient. The values of α corresponding to the various well-defined firings can be evaluated from the figure, and thus, the half-cyclic average current I_{av} is calculated as

$$I_{av}=\frac{V_g\sqrt{2}}{\pi R_L}(1+\cos\alpha)$$

where

V_g = the rms of the a-c supply voltage
R_L = the total resistance of load and gate winding
α = the angle of firing

For the case shown in Fig. 1(A), R_L is 1,200 ohms, V_g is 125 volts rms, k is 2, and the frequency is 60 cycles per second. The half-cyclic average I_{av} can also be obtained easily for each α from steady-state calibrations in which I_{av} is read on a rectified meter for various values of the α's observed over the range of calibration. Similar steady-state calibrations with an rms meter can be used to relate half-cyclic effective values I_{rms} to the angles of firing.

From the adjustment chosen it is known that traces 1, 2, 3, etc., shown in Table I, are displaced from one another by 2 cycles. A certain ambiguity remains as to the time elapsing between the application of the signal pulse and the firing recorded as $n=1$, i.e., the time elapsing between trace 0 and trace 1. This ambiguity can be reduced by the use of proper circuits which synchronize the initiation of the step with the recurrent oscilloscope sweep. The transient may also be repeated a few times at random choosing the picture which, by inspection, shows the most plausible trace number 1. Or, even better, the transient can be initiated from some initial firing angle beyond the range of interest, making use of the final results only over the range under study.

The choice of $k=2$ in this case was dictated by the recognition that the transient was slow and, therefore, the ambiguities mentioned may be rather insignificant. Of course, if the transient is fast, all positive half-cycles or even all positive and negative half-cycles can be displayed, thus directly reducing the ambiguity to the one half-cycle which is inherent to the amplifier itself.

Fig. 1(B) expresses the results of Table I. It is seen that these results greatly differ from any envelope of peaks such as the one displayed in the same figure for the same transient.

Table II. Calculation of Transient Response: Inductive Resistive Rectified Load

Trace No., n	Cycle No., kn	Firing Angle α, Degrees	$1+\cos\alpha$	i_n-i_{n-1}	I_{av}, Milliamperes
0	0	109	0.67	0	31.5
1	3	92	0.97	0	45.5
2	6	78	1.21	1.4	56.3
3	9	68	1.37	1.5	63.7
4	12	61	1.48	1.5	69.3
5	15	55	1.57	1.5	73.5
6	18	50	1.64	2.1	76.5
8	24	43	1.73	1.0	81.0
10	30	39	1.78	0.8	83.1
∞		32	1.85	0	86.5

INDUCTIVE RESISTIVE NONRECTIFIED LOADS

Here knowledge of the angles of firing α and the angles of extinction β allows the evaluation of the half-cyclic average of load current as given in eq. 19, reference 4

$$I_{av}=\frac{1}{\pi}\int_{\alpha}^{\beta} i_L d\omega t=\frac{V_g\sqrt{2}}{\pi R_L}(\cos\alpha-\cos\beta)-\frac{E_b}{V_g\sqrt{2}}(\beta-\alpha)$$

Again, as the instantaneous current extinguishes periodically in each half-cycle, steady-state calibrations can be used to relate I_{av} to α without the use of the foregoing equation. Similar calibrations can be obtained for effective values using an rms meter.

INDUCTIVE RESISTIVE RECTIFIED LOADS

As in the previous cases, with the knowledge of the angle of firing α the volt-time area of the power supply which has not been absorbed by the cores can be evaluated. However, this volt-time area is applied to the load and is shared by the load resistance R_L, the load inductance L_L, and any battery of direct-voltage source E_b (taken positive if opposing the flow of current) that may be present in the load. For continuous flow of current in the load, I_{av} is obtained from

$$V_g\sqrt{2}(1+\cos\alpha)=\pi R_L I_{av}+\pi E_b+L_L(i_n-i_{n-1})\frac{\omega}{k}$$

where i_n and i_{n-1} are the instantaneous currents at the firing instants α_n and α_{n-1} and are obtained from the picture of the transient. Hence, the term $L_L(i_n-i_{n-1})\omega/k$ can be evaluated and thus I_{av} is obtained. Fig. 2(A) shows such a transient with $k=3$. The load resistance R_L is 1,200 ohms, the inductance L_L is 0.2 henrys, the direct-voltage source is zero, the rms gate voltage is 125 volts, and the peak load current is 140 milliamperes as measured by calibrating the vertical axis of the oscilloscope. The vertical calibration is necessary to evaluate $i_n - i_{n-1}$. Steady-state calibrations do not apply here if continuous flow prevails because the half-cyclic averages of voltage absorbed by the inductance L_L are not zero during the transient. The results of the calculations tabulated in Table II are plotted in Fig. 2(B), which shows the variation of the peaks of the output current during the transient for purposes of comparison.

If the flow over a certain range is discontinuous the term $L_L(i_n - i_{n-1})\omega/k$ is zero and steady-state calibrations can be used. An equation directly relating I_{av} to α is given in the following[4]

$$V_g\sqrt{2}(1+\cos\alpha) = \pi R_L I_{av} + (\delta-\alpha)E_b$$

where δ is the angle at which the instantaneous load current extinguishes after the firing at α.

In some practical cases it is adequate to evaluate the half-cyclic areas simply by visual inspection of the screen patterns; e.g., inspection of Fig. 1 shows the trends calculated more carefully in the foregoing. In the same line of thought this visual procedure can be used, with obvious limitations, to examine transient response of amplifiers with silicon-steel cores even though no definite angles of firing can be recognized.

Frequency Response

The application of a sinusoidal signal causes the half-cyclic output of the amplifier to vary periodically at signal frequency. In many cases it may be adequate to define the response of the amplifier to a signal of given frequency and amplitude with given bias simply in terms of the maximum and minimum values of half-cyclic output. A display of output half-cyclic wave forms, as shown in Fig. 3(A), yields immediately the two limits between which the angle of firing varies periodically. From this the corresponding limits of the half-cyclic output (average or effective value) are readily evaluated as discussed in the section entitled "Transient Response."

To determine the phase shift between maximum output and crest of the applied signal, an intensity-modulating pulse derived from the signal voltage is applied to the Z-axis of the oscilloscope and is phase-shifted until it occurs during the half-cycle of maximum output. This condition is met when the firing of maximum output appears with a heavier trace on the pattern; see Fig. 4(A). The trace of the sinusoidally varying signal is intensity-modulated in the same manner and thus the phase shift ϕ between the crest of the signal and the occurrence of a maximum (or minimum) output is displayed; see Fig. 4(B). The system used is shown in the block diagram of Fig. 4(C). Phase shifts can be measured in this manner directly from the screen patterns with an accuracy of ±5 degrees. A better resolution can be obtained by the use of a calibrated phase shifter to supply the intensity-modulation pulse; in this case the phase shifts in question result from readings on the calibrated dial.

In some instances the foregoing information may not be adequate because the half-cyclic output, though varying periodically at the signal frequency, may not result in a true sinusoidal modulation envelope for a variety of well-known reasons. Therefore, it may be of interest to determine not only the maximum and minimum value of output but rather, more precisely, how the output varies throughout 1 cycle of the signal. If the gate frequency is an integer multiple of the signal frequency, stationary patterns of subsequent half-cycle output wave forms can be displayed on the screen with the same techniques as described initially. In the present case, however, the wave forms of increasing and decreasing output may appear intermingled and difficult to interpret. It is an easy matter than to apply a proper blanking voltage to the Y-axis to remedy this difficulty. This voltage is derived from the signal voltage and is phase-shifted so as to blank out the patterns of either ascending or descending output. Fig. 5 shows the response of a 400-cycle-per-second magnetic amplifier with a control frequency of 1/13 of 400 cycles per second or 30.8 cycles per second with ascending outputs blanked out.

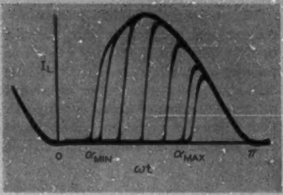
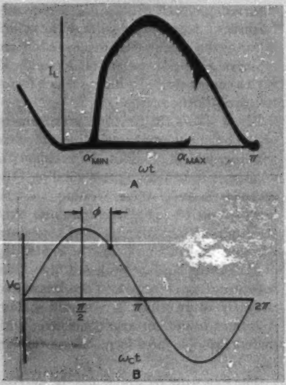


Fig. 5. Response of a 400-cycle amplifier to a 30.8-cycle signal. Traces of output throughout half-cycle of signal, k=1

Conclusions

Simple techniques have been described which allow to study experimentally the transient and frequency response of magnetic amplifiers with various types of loads (e.g., resistive or resistive-inductive with or without load rectification).



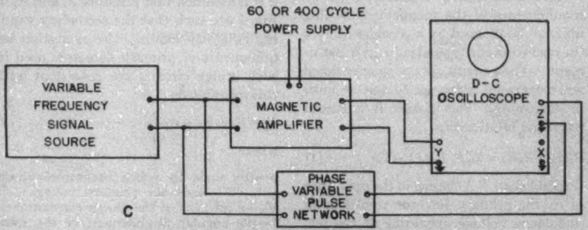


Fig. 4. Phase-shift measurement of a 400-cycle amplifier

A—Intensity-modulated output current

B—Intensity-modulated control voltage; ϕ is phase shift between peak of control voltage and maximum output current

C—Laboratory setup for measurement of phase shift

The equipment required consists of a conventional oscilloscope with the addition, in some cases, of a few other components of standard design. The information obtained may not be as accurate as can be expected from carefully designed pulse integrators, yet it is remarkably more significant that any observation of envelopes of peaks of common practice.

References

1. Magnetic Amplifier Transient Analyzer, W. A. Geyger. *Electronics*, New York, N. Y., vol. 26, no. 7, July 1953, pp. 189–91.

2. A Transient Analyzer for Magnetic Amplifiers, E. J. Smith. *AIEE Transactions*, vol. 72, pt. I, Sept. 1953, pp. 461–65.

3. The Cyclic Intergrator—A Device for Measuring the Frequency Response of Magnetic Amplifiers, T. Dunnegan, Jr., J. D. Harnden, Jr. *Ibid.*, vol. 73, pt. I, Sept. 1954, pp. 358–65.

4. The Operation of Magnetic Amplifiers with Various Types of Loads, Part I—Load Currents for Given Angle of Firing, L. A. Finzi, R. R. Jackson. *Ibid.*, July 1954, pp. 270–78.

An Angular Position A-C Electromotive-Force Transducer and Its Application to Null-Type Recorders

E. C. GUPTILL
ASSOCIATE MEMBER AIEE

C. W. McCARTY
NONMEMBER AIEE

THE measurement of process variables by means of a-c excited primary elements has been somewhat neglected. The general use of potentiometric systems has resulted largely from their excellent performance characteristics and their adaptability to a wide range of primary elements. It is generally true, however, that for a given measurement application an a-c system is simpler and hence less expensive. No provision need be made for d-c to a-c conversion, cold junction compensation, or for battery standardization.

This paper describes in detail a new element (intermediate means) which provides in one housing means for exciting a primary element and a properly phased nulling or rebalancing voltage. This is an iron core device having a primary, a secondary or excitation winding, and an air gap in which a feedback winding is servo positioned to rebalance an input error (unbalance) voltage. A previous paper[1] has described the a-c measurement system in which this magnetic balancing unit is used.

Objectives

The major objectives for the magnetic balancing unit are listed in the following, based upon its use in an instrument of the 1/4-per-cent accuracy classification.

1. The provision of an accurate relationship between shaft position and an a-c electromotive force at a given level of excitation.
2. A high degree of stability of calibration over a period of several years.
3. Reproducibility consistent with the use of predetermined scales.
4. Operation at power frequencies, eliminating the need for frequency or d-c to a-c conversion.
5. Elimination of sliding contacts with their associated wear and maintenance.
6. Suitability for manufacture in production quantities.
7. Infinite resolution to provide stepless rebalancing.

Principles of Operation

An understanding of the operational requirements of the magnetic balancing unit can be reached by a consideration of the operation of a typical industrial instrument. The circuit of a conventional servo-operated resistance bridge is shown in Fig. 1(A). This circuit is balanced when the relationship

$$R_1(R_4+R_sK)=R_2[R_3+R_s(1-K)] \qquad (1)$$

is established. A change in the resistance of R_1, the primary detector, results in an unbalance voltage appearing at the terminals of the servo amplifier. This voltage is amplified and applied to the control field of the servomotor. The servomotor drives the slidewire to a position such that equation 1 is satisfied. The circuit is again in balance. The pointer and associated end devices have been repositioned to correspond to the new value of the measured variable R_1.

It should be noted, however, that it is not necessary to rebalance the bridge itself, by proportioning the arms, to effect the proper repositioning of the end devices. Instead, circuit rebalance may be achieved advantageously by the insertion of a voltage of the proper phase and magnitude at point *A*, Fig. 1(A). This modification is shown in Fig. 1(B).

The magnetic balancing unit which is shown diagramatically in Fig. 2 not only furnishes a rebalancing voltage thus eliminating the slidewire but, in addition, replaces precision resistors R_3 and R_4 with a center-tapped winding which forms two arms of the bridge.

Fig. 3 shows the physical arrangement of the balancing unit, including the rotor winding, which provides the rebalancing voltage, the center-tapped winding (the secondary), and a third winding (the primary) used to excite the unit and thus the bridge. The primary and secondary windings are mounted on the center leg of the stator lamination assembly. The rotor winding is wound on the laminated rotor structure which is mounted so that it rotates about an axis concentric with cylindrical pole faces formed in the center leg.

The following relations may be derived if it is assumed that resistors R_1 and R_2 in Fig. 2 are such that the secondary winding is lightly loaded; the excitation requirements of primary detectors used in such bridge circuits are consistent with this assumption

$$V_r=K_rV_p\theta\angle\phi \qquad (2)$$

where

ϕ = the angle by which the rotor voltage V_r leads the primary voltage
V_p = a function of the design constants
θ = the angular displacement of the rotor from its neutral (reference) position
K_r = a design constant

In order that the rotor voltage may rebalance the bridge circuit, it is necessary that the bridge excitation be of the same phase as the rotor voltage. Since the

---

Paper 55-547, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 31, 1955; made available for printing May 9, 1955.

E. C. Guptill and C. W. McCarty are with the General Electric Company, West Lynn, Mass.

Vol.
Pt. 1
1955

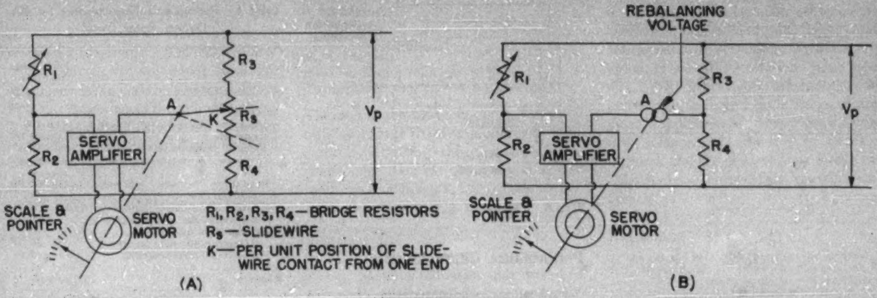

REBALANCING VOLTAGE
R1
R2
R3
R4
Rs
A
K
Vp
SERVO AMPLIFIER
SCALE & POINTER
SERVO MOTOR
R1, R2, R3, R4—BRIDGE RESISTORS
Rs—SLIDEWIRE
K—PER UNIT POSITION OF SLIDE-WIRE CONTACT FROM ONE END
(A)
(B)

Fig. 1. Servo-operated resistance bridge

A—Conventional B—Modified

introduction of the air gap in the magnetic circuit results in a rotor voltage which leads the primary voltage by a considerable angle, depending upon the resistance of the primary winding and the dimensions of the gap, the phase of the bridge excitation voltage must be similarly displaced. Obviously the secondary winding, linked by the same flux as is the rotor winding (and by leakage flux of the same phase) and lightly loaded, provides a properly phased excitation source.

The secondary voltage V_s is expressed as

$$V_s = K_s V_p \angle \phi \tag{3}$$

where K_s is a design constant.

The approximate equivalent circuit of Fig. 4(A) and the phasor diagram, Fig. 4(B), indicate these relationships. It should be noted that these figures represent the conditions that exist at one rotor position only. The tap on the mutual reactance X_m indicates that part of the flux linking the secondary does not link the rotor winding.

The equation for the unbalance voltage of the bridge is

$$V_b = \frac{V_s}{2}\left(\frac{R_2-R_1}{R_2+R_1}\right) \tag{4}$$

When the rotor voltage V_r is equal in magnitude to V_b and oppositely phased, null balance obtains, corresponding to the rotor position

$$\theta = \frac{K_s}{2K_r}\left(\frac{R_2-R_1}{R_2+R_1}\right) \tag{5}$$

It should be apparent that, although a simple resistance bridge circuit has been used to explain the operating principles of the magnetic balancing unit, its usage is not confined to this application. For example, resistance strain gauges of the bridge type, which are widely used in the measurement of pressure, flow, liquid level, etc., can be excited from the balancing unit secondary, their output balanced against the rotor voltage. In this case, the rotor output may be modified through the use of an additional winding or by mechanical means to correspond to the output voltage of strain gauges which usually are of one polarity, i.e., vary from zero to maximum as the measured variable covers its range. As another example, differential transformers excited in series with the balancing unit to obtain proper output voltage phasing may be usefully employed as primary detectors in the measurement of displacement, pressure, etc.

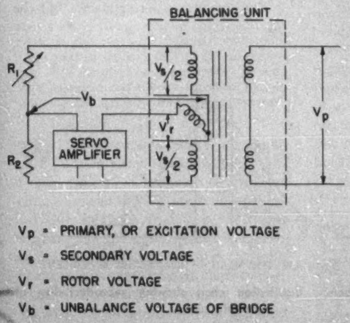

BALANCING UNIT
R1
R2
Vb
Vs/2
Vr
Vs/2
Vp
SERVO AMPLIFIER
Vp = PRIMARY, OR EXCITATION VOLTAGE
Vs = SECONDARY VOLTAGE
Vr = ROTOR VOLTAGE
Vb = UNBALANCE VOLTAGE OF BRIDGE

Fig. 2 (left). Diagrammatic sketch of magnetic balancing unit

Construction

In Fig. 5 some details of the construction of the magnetic balancing unit can be seen. The stator and rotor structures are made up of annealed silicon-steel laminations. The stator laminations are circular, for improved mechanical stability, and are rigidly mounted on a heat-treated cast aluminum frame which also houses the rotor bearings. In assembly the stator is located by means of a gauge inserted in the rotor bearings to insure a high degree of reproducibility between units.

The rotor lamination assembly is staked, centerless ground and assembled concentric with the shaft, which is also centerless ground. Spring-loaded ball bearings aid in reducing side play and end play to a minimum. The effects of end play are further reduced by making the rotor stack height greater than the stator stack height.

The excitation and secondary windings are wound on spools which are mounted on the poles of the stator lamination, as shown in Fig. 3. To help insure a precise

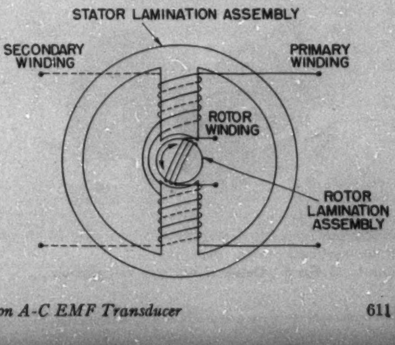

STATOR LAMINATION ASSEMBLY
SECONDARY WINDING
PRIMARY WINDING
ROTOR WINDING
ROTOR LAMINATION ASSEMBLY

Fig. 3 (right). Elements of the magnetic balancing unit

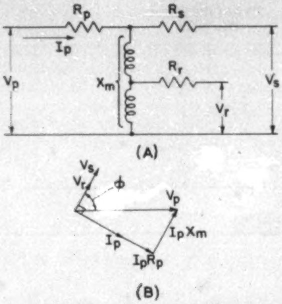


Fig. 4. (A) Approximate equivalent circuit of balancing unit, secondary and rotor windings not loaded. (B) Phasor diagram of balancing unit

division of the secondary voltage on each side of the point of attachment of the feedback winding, each arm of the secondary is made up of more than one individual winding, as indicated in Fig. 6. Successively connected windings are located on opposite legs of the lamination assembly. In addition, as an aid in achieving product uniformity, each winding consists of one strand of an assembly of twisted strands, each conductor insulated from the others. Thus, all strands occupy the same average position in a coil, so that equal voltages are induced in the windings.

To minimize the effects of stray magnetic fields, the unit is enclosed in an annealed steel can. Three slotted mounting pads are provided in the cast aluminum base. The span adjustment control, in the form of a wire-wound potentiometer across the rotor winding, is accessible through a hole in the end of the shield can.

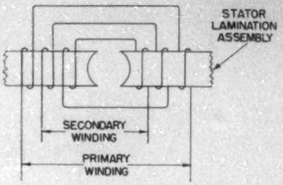


Fig. 6. Details of balancing unit windings

Performance Characteristics

Performance characteristics of representative balancing units are summarized in Table I. These data are based on 60-degree rotor motion. Improved linearity is obtained by reducing the full-scale rotor movement. Thus, for a 50-degree travel, the departure from linearity is less than 0.05 per cent, and for 40-degree travel, less than 0.02 per cent. Obviously some sacrifice of stability is to be expected as the full-scale angle is reduced, inasmuch as mechanical changes become a greater percentage of the full-scale travel.

The high degree of linearity demonstrated by the balancing unit is achieved by close control of the factors determining the uniformity of air-gap flux density, such as gap geometry, magnetomotive-force drop in the pole faces, leakage flux, etc. A constant air-gap flux density results in linearity of rotor voltage versus angular position since, under this condition, a given increment of angle always results in a given change in rotor voltage regardless of the initial angular position.

As is indicated in Table I, ambient temperature influence depends on the circuit in which the balancing unit is used. For a typical bridge circuit as shown in Fig. 2 the effect is associated with the change of resistance of the secondary winding expressed in terms of the total circuit resistance. By properly loading the rotor winding, it is possible to nullify this effect and, if it is desirable, to compensate for temperature influence in other parts of the circuit. Generally speaking, it is quite practical to limit ambient temperature influence to 0.1 per cent or better over the range of −20 to +50 degrees centigrade. In those instances in which the secondary winding is not used to excite the transmitting element, temperature influence is minimized by maintaining a high resistance in the span adjustment circuit. The performance of the balancing unit is not influenced by changes in excitation voltage and frequency when used in circuits similar to that shown in Fig. 2 or when the transmitting element and the balancing unit are excited from the same source.

Table I. Performance Characteristics for 60-Degree Rotor Travel

	Per-Cent Full Scale
Accuracy	
Linearity	±0.10
Hysteresis	less than 0.01
Ambient temperature influence, −20 to +50 degrees centigrade	depends on application — see text
Excitation voltage influence, ±10 per cent	less than 0.01
Excitation frequency influence, ±5 cycles per second	less than 0.01
External field influence	
5 oersteds, 60 cycles per second	±0.05
5 oersteds, direct current	less than 0.01
Stability	±0.05
Reproducibility	±0.03

Because of the great importance of stability in establishing the performance level of an instrument, an extensive life test program was carried out. Pilot plant production units have demonstrated excellent stability (±0.05 per cent) during accelerated tests. Units were temperature-cycled over a 6-month period from −20 to +50 degrees centigrade, with 2 complete cycles per week.

The high degree of reproducibility to be

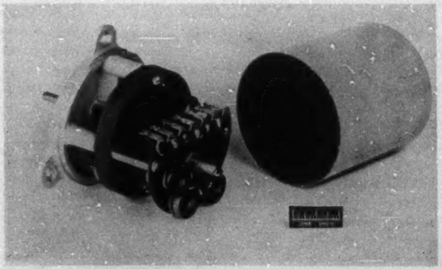


Fig. 5. Details of balancing unit assembly

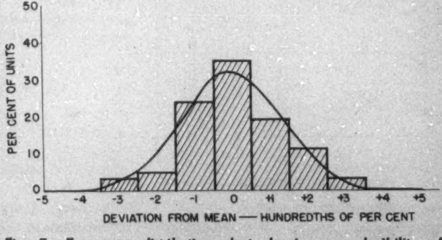


Fig. 7. Frequency distribution chart showing reproducibility of balancing units

expected is demonstrated in Fig. 7. The chimney chart presents data taken on 18 representative units. The relative frequency curve was computed from the arithmetic mean and the standard deviation corresponding to the chimney chart. It is apparent from the relative frequency curve that substantially all units will fall within limits of ±0.03 per cent from the mean.

The design constants for 60-cycle-per-second operation of the balancing unit are:

V_p, 3.2 volts.
V_s, 4.2 volts.
I_p, 0.90 ampere.
ϕ, 45 degrees.
R_p, 2.5 ohms.
R_s, 5.0 ohms (total).
R_r, 1.8 ohms.
$\Delta V_r/\Delta\theta$, 2 millivolts per degree.
Temperature rise, 15 degrees centigrade.

The use of the balancing unit is not restricted to this frequency, however. By suitably adjusting the excitation level, proper operation can be obtained at frequencies as high as 800 to 1,000 cycles per second.

Conclusions

The magnetic balancing unit described provides an accurate, stable shaft position versus a-c electromotive force relationship. The further objectives, elimination of sliding contacts and the provision of infinite resolution, have been met. Performance data indicate that the unit is usable over a wide range of ambient conditions of temperature, stray field, voltage, and frequency, in the measurement and the control of process variables such as temperature, pressure, and flow.

Reference

1. ADVANCES IN NULL-TYPE RECORDER DESIGN, **I. F. Kinnard, M. A. Princi, A. Hansen.** Paper no. *54-IRD-12*, American Society of Mechanical Engineers, New York, N. Y., 1954.

---

l. 74
1
55

Frequency-Shift Signaling Circuit for 45-Type Carrier Systems

K. E. APPERT
NONMEMBER AIEE

R. S. CARUTHERS
MEMBER AIEE

Synopsis: Signaling equipment used with short-haul carrier systems is now supplied largely on an out-of-band basis. On-hook, off-hook signals are obtained either through turning a single tone on and off, or through shifting the tone alternately between two frequency positions. Signaling with Lenkurt 45-type carrier systems makes use of frequency-shift (FS) methods. Laboratory tests with artificially induced noise and field tests during lightning periods show at least a 6-decibel (db) advantage of FS over on-off signaling arrangements in freedom from pulse distortion effects. FS arrangements also show little impairment in performance over all expected operating conditions in a telephone office. The FS signaling arrangement described in this paper also provides automatic channel level regulation of the speech channel through availability of the signaling tone as an ever-present regulating tone. Varistor keying of a transistor oscillator is used in the transmitting terminal and a polar relay in the receiving terminal to supply all normal varieties of *M* and *E* lead signaling. Mechanical arrangements provide very compact assemblies through plug-in subunits and utilization of printed wiring.

WITH the rapid growth of communication networks in recent years, interconnection of a number of telephone circuit links in tandem are frequently involved in establishing a call between subscribers. In the past, these interconnections have involved circuits operating mostly at voice frequency in short connections, and at carrier frequency only on the long toll connections. For the most part, the short-haul links have used tone methods only for ringing purposes and d-c methods for dialing. With carrier now beginning to dominate even the short-haul field and, in some instances (e.g., with rural carrier) spreading into the subscriber links, tone dialing and signaling frequently are employed in each link of the subscriber interconnection.

Short-haul carrier connections make use almost exclusively of out-of-band signaling. In contrast, longer circuit connections ranging to transcontinental distances have used in-band signaling, with the tone turned on while making connections and turned off when the subscribers were connected. Interconnections of various combinations of short- and long-haul circuits may involve both out-of-band and in-band signaling, with further varieties in the particular type of in-band and out-of-band signaling employed. Out-of-band signaling is either of an amplitude-modulated (AM) tone-off, tone-on variety or a FS type between two out-of-band tones to provide the on-hook, off-hook signal. There has been no well-established or orderly procedure in design of signaling systems to select in-band versus out-of-band signaling, or on-off versus FS signaling. Variations appear in the arrangements provided by manufacturers in this and other countries without clear presentation of the reasons for choice of the signaling methods employed.

The choice of out-of-band signaling has been based on the belief that costs are lower and circuits simpler with this method of operation. However, it has been generally assumed that this method of signaling was limited to use in one or two short links in tandem. Most of the limitation was supposed to come from the additional receiver and transmitter circuit used at each connecting point. With improved signaling performance being incorporated into the design of out-of-band signaling arrangements, however, there is probably little justification for the fears of signaling impairment introduced through multilink connections. Moreover, as short-haul carrier with out-of-band signaling is utilized at longer and longer distances, any performance disadvantages of out-of-band signaling compared with in-band signaling become relatively unimportant. Already, use of such carrier has been extended to distances frequently as great as 200 to 300 miles. The tendency toward higher costs of in-band signaling comes chiefly from the need for special networks and switching means to block the in-band tone from the subscriber circuit during certain periods of call completion or in no-charge telephone calls.

For a number of years, out-of-band signaling has employed on-off signaling rather than FS signaling in the belief that it was more economical and required less frequency space. Designers, at the same time employing methods of auto-

---

Paper 55-560, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 29, 1955; made available for printing May 3, 1955.

K. E. APPERT and R. S. CARUTHERS are with the Lenkurt Electric Company, Inc., San Carlos, Calif.

Table I. Pulse Distortion Limits Used in Design of Lenkurt Carrier Systems

	Signaling Pulse Distortion at Normal Battery and System Levels, Per-Cent Break
Dial	63.5 ± 4
Outgoing trunk	−6 ± 2
First carrier link	+1 ± 2
Pulse link	−1 ± 1.5
Second carrier link	+1 ± 2
	58.5 ± 11.5

Table II. Extremes of Pulse Distortion in Toll Connection

	Signaling Pulse Distortion, 3γ-Per-Cent Break
Dial	63.5 ± 4
Outgoing trunk	−6 ± 2
First carrier link	+1 ± $3\sigma_c$
Pulse link	−1 ± 1.5
Second carrier link	+1 ± $3\sigma_c$

Summation: $(3\sigma) = 58.5 \pm \sqrt{(3\sigma_1)^2 + (3\sigma_2)^2 + \ldots + 2(3\sigma_c)^2}$

Table III. Pulse Distortion During Lightning Storm

Noise on AM Carrier Channel Measured at Expandor Output, Decibels Adjusted at 0 Level	Noise Computed at Expandor Input, Decibels Adjusted at 0 Level		Pulse Distortion, Per-Cent Break	
	AM Channel	FS Channel*	AM Channel	FS Channel
≫50	≫68.5	≫60.5	5	10
≫50	≫68.5	≫60.5	failure	failure
45–50	66–68.5	58–60.5	4	1
45–50	66–68.5	58–60.5	2	<0.5
45–50	66–68.5	58–60.5	2	<0.5
45–50	66–68.5	58–60.5	4	<0.5
45–50	66–68.5	58–60.5	2	<0.5

* Assumes 8 db higher operating level and 8 db less noise than on carrier channel with AM signaling.

Table IV. Effects of Steady Noise on Pulse Distortion

Flat Noise (0 Message Level)						
Pulse Distortion, Per-Cent Break	At Expandor Output, Decibels Adjusted	Compandor Noise Advantage, Decibels Adjusted	At Expandor Input, Decibels Adjusted	Dbm		Signal-to-Noise Radio, Db
				2,750-Cycle Band	150-Cycle Band	
FS Signaling (Signal Power—16 Dbm at 0 Level)						
1	35	26	61	−21	−33.5	17.5
2	40	23.5	63.5	−18.5	−31	15
4	47.5	19.5	67.0	−15	−27.5	11.5
5	50	18.5	68.5	−13.5	−26	10
10	55	16	71	−11	−23.5	7.5
15	57.5	14.5	72	−10	−22.5	6.5
AM Signaling (Signal Power 0 Dbm at 0 Message Level)						
1	50	18.5	68.5	−13.5	−26	26
2	60	13.5	73.5	− 8.5	−21	21
4	67.5	9.5	77	− 5	−17.5	17.5
5	70	8.5	78.5	− 3.5	−16	16
10	73	7	80	− 2	−14.5	14.5
15	74.5	6	80.5	− 1.5	−14	14

matic channel level adjustment to keep signaling performance within tolerable limits, attempted to take advantage of signaling tones to provide regulating information. Various types of memory circuits were considered to maintain levels during the tone-off period, with little practical success. In some instances transmitted carriers were used in the channels as a means of level adjustment. More recently, Lenkurt short-haul carrier systems with suppressed carrier have utilized FS signaling tones always present and suitable for regulation purposes. Both the FS and tone-on, tone-off signaling arrangements typical in present-day short-haul carrier systems use about the same bandwidth. FS channels show considerable improvements in performance over on-off signaling channels under typical operating conditions. The evidence is that FS channels generally could be operated with narrower bands and lower levels with performance equal to on-off signaling. Even without inclusion of such refinements as limiters and delay circuits frequently made a part of FS circuits, the balancing action of the discriminator detector circuit and polar relay against noise is of great benefit.

Performance of the FS signaling channel used in the Lenkurt 45-type carrier equipment is described herein. Performance is shown as measured under varied operating conditions in the laboratory and under field operating conditions. Comparisons have been made also with on-off signaling circuits operating under similar conditions. These include exposure to high noise during severe lightning storm periods.

Signaling Requirements

Much of the carrier equipment now being designed is being operated as intertoll trunks terminating in telephone offices using step-by-step switching equipment. A generally accepted limit used in design is that the total distortion of pulses operating through the various trunks between the subscriber-end step-by-step office should range between about a 44- and 72-per-cent break, with good performance in terms of producing minimum wrong numbers from missed dial pulses. For example, at 10 pulses per second, a 60-per-cent break indicates on-hook for 60 milliseconds and off-hook for 40 milliseconds. In the design of Lenkurt type-45 carrier systems,[1] distortion permitted in the parts of the over-all trunk was considered to be distributed as shown in Table I.

During extreme battery and level conditions, each carrier link was expected to have about ±2 per cent more pulse distortion than during normal operating periods, and just barely fall within the step-by-step office pulse distortion limits. In these performance objectives, design of the carrier signaling channel was rather severely taxed. Such a performance requirement, however, is probably more severe than is warranted, because the distortion contributions from the various parts of the over-all connection were not considered as additive on a probability basis. If the distortion permitted in each part of the circuit connection is considered to be the 3-sigma (3σ) point on a normal distribution curve, then total permissible distortion can be computed as shown in Table II. If the total distortion 3σ is permitted to be ±13.5 per cent, then the carrier channel distortion can be $\sigma_c = \pm 3$ per cent, and $3\sigma_c = \pm 9$ per cent.

Considerably better performance than this occurs in the signaling channel of 45-type carrier equipment during normal periods of time. These wider limits of permissible per-cent break distortion are of interest only during periods of high lightning activity. It seems certain that distortion produced by lightning would not simultaneously reach large magnitudes in both carrier links of the example. If during periods of maximum distortion (3σ) in the over-all link only one carrier link is contributing its worst distortion, then the lightning-disturbed link might be permitted to reach a distortion of $3\sigma = \pm 12^1/_2$ per cent. This indicates that in 1,000 cases where lightning is causing pulsing distortion in the carrier equipment, only two cases occur with distortion exceeding $12^1/_2$ per cent, and thus probably causing wrong operation of dial equipment. There is danger, of course, in using calculations of this type without considerable interpretation. A more reasonable conclusion is probably that when pulsing distortion from impulse disturbances does occur, little damage will be done unless distortion exceeds 10 per cent. The chief danger is in having distortion attain such high levels for long periods of time.

High levels of lightning disturbance are evidenced by high levels of noise in the voice channels of the carrier system. Little interference is likely in the signaling channel with such high noise bursts unless the burst is coincident with the transition interval between on-hook and-off-hook whereby pulses would be shortened or lengthened. In a normal telephone call with an average connected time of 3 minutes, dialing time takes about 10 seconds. About 35 dial pulses are sent in a 7-digit call. Since 100 10-pulse-per-second dial pulses might be sent in 10 seconds, there are only 35/100 or about one-third chance that dial pulses are being sent at the time of interference during the 10-second dialing period. If, for example, during the worst 10-minute period (with 60 possible dialing periods) of a severe lightning storm, it were observed that 20 pulses of noise caused dial pulse distortion exceeding 10 per cent, the probability of causing wrong numbers would be: $20 \times 1/3 \times 1/60 = 1/9$, or 1 in 9 calls.

Operating Levels for FS and AM Signaling

As more and more experience is gained in use of dial signaling in carrier telephone equipment, it becomes evident that various interferences, either from other office equipment such as relays, from speech, or from natural disturbances such as lightning, are the factors of greatest concern in design. The older types of carrier systems used on the heavy toll trunk routes have been able to bear the expense of careful engineering to guard against noise disturbances. Higher operating levels and more expensive measures have been utilized in the signaling channels to guard against noise. Newer carrier systems have been applied in existing cable and open-wire facilities where carrier use was not intended and where the expense of improving facilities could not be justified. Furthermore, these systems for economic reasons have been stretched in using both lower operating levels and longer spans. To protect the speech channels against noise during normal and heavy static periods including lightning, a compandor has often been associated with each message channel. The compandor affords noise advantage only to the message channels and gives no basic advantage to signaling performance. Adequate signaling performance is obtained under such adverse conditions either through materially increasing the signaling levels or by incorporating certain noise-guarding devices or both. With AM signaling, it is common practice to operate the tone-on signaling magnitude at 0 decibels referred to 1 milliwatt (dbm) at the 0 message level point. The older, longer-haul carrier systems, on the other hand, use the AM signaling tone at a magnitude of −16 dbm or lower, at the 0 message level. In addition, other artifices are employed such as delay networks in the signaling receiver to guard against the impulse nature of noise disturbances. This further complicates circuitry by requiring special level-compensating circuits in the receiver. Variations in levels of received signals otherwise would result in excessive signaling distortion from the distorting effects of the rise and decay times of the delay network

Even the older carrier systems utilized on well-engineered outside plant facilities have resorted to special artifices in addition to higher operating levels. To guard against noise interference to dialing, higher signaling levels are utilized during dialing through the incorporation of slow-acting relays that are not actuated during dial trains. To protect supervisory signal operations from speech interferences, guard circuits with additional relays are sometimes employed. This latter diffi-

Fig. 1 (right). Effects of noise on signaling distortion

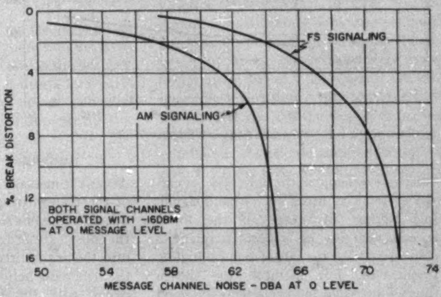
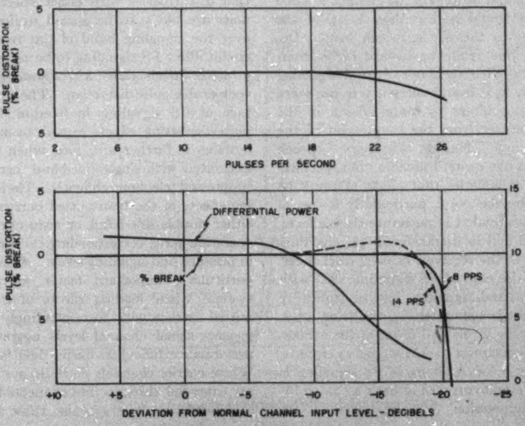


Fig. 2 (below). Effects of pulsing speed and signal level on distortion



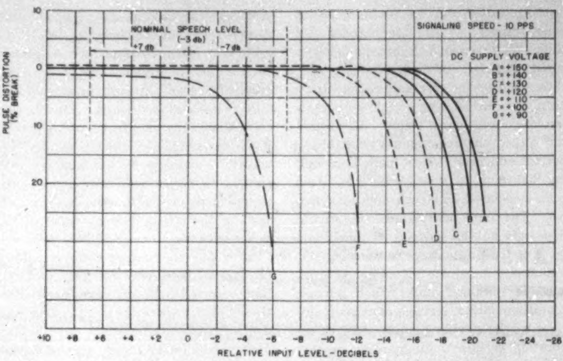


Fig. 3 (above). Effects of plate voltage variations on signaling distortion

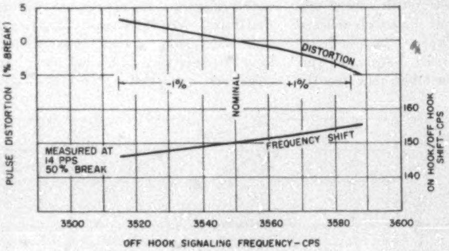


Fig. 4 (left). Effects of frequency error on signaling distortion

culty is overcome in the new, cheaper carrier systems by using out-of-band signaling, taking advantage of the lower speech energies of speech frequencies coincident with the out-of-band signaling frequencies. For example, the speech interferences in a 150-cycle signaling band that might be used in the vicinity of 3,500 cycles would be less than 1/100 of the energy in the whole speech band. Despite this tendency toward lower interference with such out-of-band signaling channels, it is still necessary to use voice low-pass filters to block speech in the transmitter from the frequencies of the signaling channel. Ordinary speech would be quite tolerable, but peculiar lisps or whistles that might at times be sent by the voice, particularly if special effort is made to exaggerate these effects, would end in a transfer of the signaling relay in the receiver, causing disconnect.

In the case of FS signaling used with out-of-band signaling, less susceptibility to noise is encountered, permitting lower operating levels and fewer of the attachments common to AM signaling circuits. Where 3,400/3,550-cycle FS signaling is used, employment of as little as 10- to 15-db suppression (signaling power −16 dbm at 0 message level) across this part of the transmitting speech band is adequate to prevent relay transfer even with the most extraordinary of speech sounds into the transmitter. A combination of the balancing action of the discriminator circuit with FS signaling and the fact that disturbances with either speech or noise are likely to be spread uniformly over the signaling band of the receiver circuit allow FS signaling to be operated at lower levels than AM signaling for comparable pulse distortion. The advantage of FS signaling in freedom from pulse distorting effects caused by interferences is further amplified when it is associated with single-side-band, carrier-suppressed telephone channels. The loading effects of the transmitted carriers in either double-side-band or twin-channel operation force corresponding reductions in message and signaling levels. This is a particularly important factor in radio systems where loading effects of transmitted carriers with correspondingly low message-signal channel levels aggravate performance difficulties during deep fades. Where carrier channels on radio systems are extended through interconnection to cable and open-wire systems, these difficulties become cumulative. The benefits are large in these cases from FS signaling and higher operating levels in single-side-band transmission.

In carrier systems of the Lenkurt *45BN* type used in existing single-toll type of cables, 24 single-side-band channels are obtained. Frequency bands range from 40 to 140 kc for one direction of transmission and from 164 to 264 kc in the opposite direction. Suppressed carrier operation and FS signaling are employed. Despite the low transmitting levels used and losses of as much as 64 db at the top ends of the transmitted groups, performance reaches high standards even during heavy interference periods.

Effects of Noise on FS/AM Signaling

In a field installation of the *45BN* type, comparative tests were made of AM versus FS signaling when subjected to artificially induced noise of a steady nature, and when subjected to heavy interferences during a lightning storm. The FS signaling channel would normally be operated at −16 dbm at the 0 message level point and the AM channel operated at 0 dbm at the 0 message level point, a 16-db level difference. It was expected, however, that the FS channel would show under such operating conditions a performance disadvantage of only about 8 db compared with AM signaling in freedom pulse distortions coming from lightning disturbances. To offset this, the FS signaling channel was operated at 8-db higher signaling and message levels on the lines than the AM message and signaling channels. Ten pulse-per-second signals at 50 per-cent break were sent from the transmitting terminals, and a dual-beam oscilloscope was used for simultaneous presentation of pulse distortion in the AM and FS signaling channels. In each case the signaling channel bandwidth was about 150 cycles wide. Results of tests during one heavy storm period are shown in Table III. A Western Electric *2B* noise meter was used to read noise magnitudes in the message channel and, consequently, noise peaks from heavy impulses could be read only approximately. Noise was read on the channel with AM signaling and estimated for the channel with FS signaling, making allowance for the 8-db difference in operating level. In each case the expandor part of the compandor was used in the receiving circuit, thus accounting for the high levels of noise indicated at the expandor input.

Table IV shows comparisons of AM and FS signaling pulse distortion when subjected to steady noise of high magnitudes

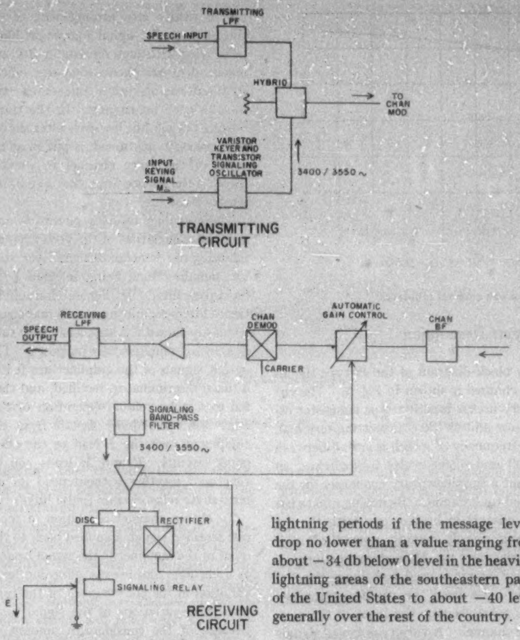


Fig. 5. Block diagram of signaling arrangements

in laboratory tests. AM signaling was operated at 0 dbm at 0 message level, and FS signaling at −16 dbm at 0 level. To compare signal-to-noise ratio for equal pulse distortion, correction was made in the *2B* noise meter readings, assuming a 2,750-cycle message band and a 150-cycle band in the signaling channel.

In Fig. 1 pulse distortion plots against noise are shown, calculated from Table IV, for AM and FS signaling, each assumed operated at −16 dbm at 0 message level. FS signaling appears to have at least 6-db advantage over the AM signaling.

Signaling channels in carrier systems operated on open-wire lines are made less subject to disturbances from lightning because of the care that is exercised in use of higher operating levels. In contrast, cable carrier systems of the Lenkurt *45BN* type operate satisfactorily on aerial cables with message levels as low as 60 to 70 db below the toll test-board level (−8 dbm of signaling power is used with FS signaling at 0 message level). On open-wire lines, comparable satisfactory signaling operation will be obtained during heavy lightning periods if the message levels drop no lower than a value ranging from about −34 db below 0 level in the heaviest lightning areas of the southeastern parts of the United States to about −40 level generally over the rest of the country.

Performance of FS Signaling Channels

Information on signaling channels during normal periods of time in the absence of noise is usually portrayed by effects of input level variation on signal distortion, effects of signaling with varying pulse rates representing normal speed in dialing (10 pulses per second considered average and 8 to 14 pulses per second extremes), and effects of battery voltage variations. These data for the signaling unit of the 45-system are given in Figs. 2 and 3. The curves show that the FS signaling channels in Lenkurt *45*-type equipment would continue to operate satisfactorily even under conditions where the message channels had become inoperative. The channels of *45*-type carrier were designed for little more than ±7-db variation from normal in receiving channel speech levels. No effects on signaling occur from plate battery voltage variations over ranges expected in telephone office use, including extremes during a-c power failure periods.

A displacement of transmitted signaling frequencies with respect to the band of the receiving channel signal filter and discriminator results in pulse distortion. The top curve of Fig. 4 shows that a 1-per-cent shift results in about a 4-per-cent break pulse distortion in a *45*-type carrier. When a single oscillator is keyed in FS signaling, both the on-hook and off-hook signal frequencies, if in error, are likely to err by about the same amount; see lower curve, Fig. 4. In general, the frequencies of signaling oscillators, the carrier frequencies, and the tuning of the receiving signal channel band filter and discriminator are of adequate stability to reduce pulsing-error distortion effects to a fraction of 1 per cent.

In FS signaling, the effect of keying separate oscillators versus that of continuously moving the frequency of a single oscillator is an important consideration in obtaining minimum pulse distortion.

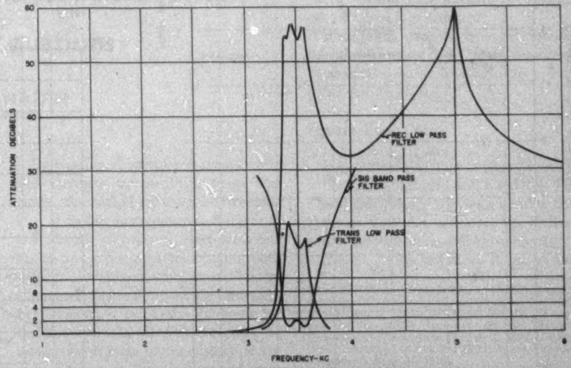


Fig. 6. Transmission characteristics of signaling filters

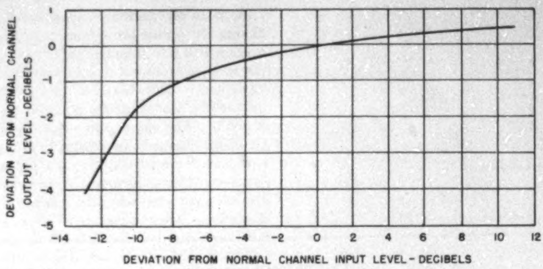


Fig. 7. Level stabilization of automatic channel regulator

The difficulty with keying separate oscillators on and off is that the bandwidth of the signal receiving filter has enough decay and build-up time to make a factor of uncertainty in the release and operate time of the signal relay in the receiving circuit. In the case of the *45*-type signaling channel with a bandwidth of about 150 cycles and a delay time near 7 milliseconds, the uncertainty in release or operate time of the signaling relay may be as much as 5 or 6 milliseconds. This represents generally unacceptable distortion for normal periods of use, and makes it desirable to use a single oscillator keyed between the two frequencies used for on-hook and off-hook. This results in elimination of this source of dial pulse error.

Circuit Description

A block diagram of the *45*-type signaling channel is shown in Fig. 5. The circuitry uses a junction-type transistor oscillator unit at the transmitting terminal, the frequency of which is varied between 3,400 and 3,550 cycles by switching in or out a supplementary condenser for the tuned tank circuit. Switching makes use of varistors in bridge form with the d-c keying signals coming from the trunk circuits over an *M* lead connection. A bridge circuit is used to balance the dial pulse noise from the oscillator circuitry and, subsequently, the input of the message channel. A variety of keyed signals can be accepted on the *M* lead, permitting use of nearly any arrangement of d-c ringdown or dial signal equipment likely to be used with such circuits in Bell and independent telephone company offices. In ringdown operation, one of the two tones is used for ringing. In the transmitting branch, the low-pass filter shown, as previously mentioned, is put in at the input of the voice channel to prevent speech interference into the signal frequency band.

The signaling receiver circuit is connected to the output of the voice channel following the demodulator amplifier, with the signaling tone being selected by a band-pass filter. In Fig. 6, characteristics of filters separating speech and signal bands are shown. A supplementary tube is used for amplification purposes. The output signals of the amplifier are fed to a tuned discriminator, rectified, and then fed to a polar relay. Operation of this relay sends keyed d-c signals from the equipment over the *E* lead to the office trunk circuits. *E* and *M* leads can be connected together without need for intermediate relay circuits (pulse links).

To obtain channel regulation, a separate rectified signal is applied back to the input of the amplifier in the signal branch as a variable bias means, affording large variation in the plate current of the tube with level variations of the signal tone. Because of the considerable amount of feedback used in this amplifier circuit, its gain to signaling tones is unaffected by d-c bias variations. The plate current variation, in turn, is caused to flow through a thermistor in bridge form at the input of the channel circuit preceding the demodulator, and thus serves as an automatic channel regulation means. Level stabilization afforded by the regulator is shown in Fig. 7. To cause any appreciable change in the level of the

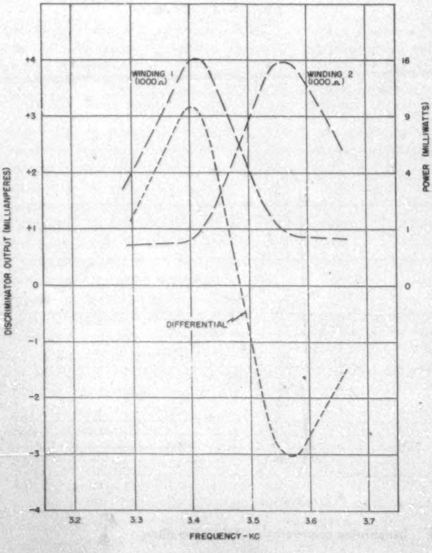
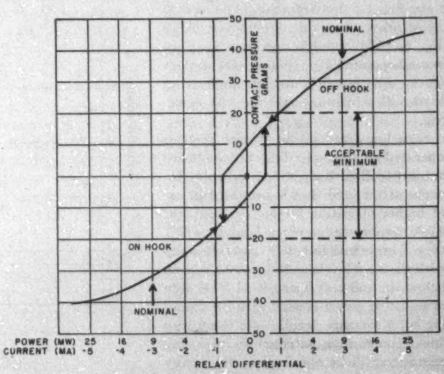


Fig. 8 (left). Discriminator characteristics

Fig. 9 (below). Contact-pressure characteristic of signaling relay



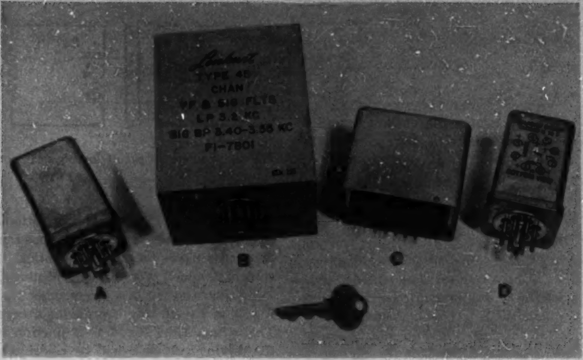


Fig. 10. Signal circuit units

A—Discriminator
B—Filter separating speech and signal bands
C—FS oscillator
D—Polar relay

signaling tone reaching the discriminator, received signal level variations in the channel must exceed ±7 db.

An important consideration in the design of such a signaling receiver is that adequate d-c power be supplied to the relay for positive contact pressure. Curves showing performance of the receiver circuit and relay used in the *45*-type signaling unit are shown in Figs. 8 and 9. The polar relay used is of good sensitivity and reliability, tests indicating that chatter-free operation can be expected over many years of use. Although a differential power is used in the windings (of about 10 milliwatts), chatter-free operation occurs with any power in excess of 1 milliwatt and transfer at 1/2 milliwatt. Normally, contact pressure is about 30 grams, chatter starting somewhat below 20 grams (1-milliwatt power differential).

Mechanical Features

The mechanical arrangements used with the signaling equipment of the *45*-type carrier take advantage of the use of a transistor. The heat problem is virtually nil, and freedom from maintenance needs in field operation with transistors justifies burying of the transistor units within the heart of the equipment assemblies. Evidence of this is that such units have been experimentally in operation in telephone system circuits since May 1954, with no maintenance problems on the signaling assemblies. Printed wiring is used throughout in the mechanical arrangements, allowing very compact assemblies and greatly benefiting factory construction where solder dipping methods are also brought into play.

Signaling unit subassemblies and the filter assembly used to separate speech and signal bands are shown in Fig. 10. These units plug in and, with the exception of the filter assembly, are removed when the out-of-band signal channel is not required. The signal receiver-amplifier circuit, also used for channel regulation purposes, is not shown as it is included in one of the main print-wired cards of the channel unit assembly.

References

1. APPLICATION AND TRANSMISSION FEATURES OF A NEW 12-CHANNEL OPEN-WIRE CARRIER SYSTEM, K. E. Appert, R. S. Caruthers, W. S. Chaskin. *AIEE Transactions*, vol. 73, pt. I, Mar. 1954, pp. 18–27.

2. THE TYPE N-1 TELEPHONE CARRIER SYSTEM, R. S. Caruthers. *Bell System Technical Journal*, New York, N. Y., Jan. 1951.

3. A SIGNALING CIRCUIT FOR N-1 CARRIER, L. A. Weber. *Bell Laboratories Record*, New York, N. Y., Apr. 1953.

4. A NEW SINGLE FREQUENCY SIGNALING SYSTEM, A. Weaver. *Ibid.*, Feb. 1954.

Formulation of Magnetic Amplifier Circuit Equations

D. H. SCHAEFER
ASSOCIATE MEMBER AIEE

MODERN magnetic amplifier circuitry makes use of two highly nonlinear elements, namely, rectifiers and square-loop magnetic cores. This paper presents a systematic method of analyzing circuits containing any sort of piecewise linear element that does not involve energy storage. The analysis utilizes a unique method of formulating circuit equations and a new type of symbolism and algebra, especially developed for the purpose, that allows operations on and simplifications of the resultant expressions.

By treating square-loop magnetic cores and rectifiers as piecewise linear elements the method provides a systematic way of finding instantaneous voltages in any circuit of resistors, rectifiers, and square-loop cores, given the instantaneous values of the driving voltages and knowledge of whether or not cores are saturated. Furthermore, it is possible to determine the times of saturation of the cores by integration if some of the history of the cores is known. The rectifiers and cores are only somewhat idealized as the effects of magnetizing currents, eddy currents in the core, and back leakage of rectifiers can be taken into account.

Ideas and symbols from a previous paper[1] will be used here. An attempt will be made, however, to make the present paper self-contained.

Square-Loop Cores

If a sinusoidal voltage is placed in series with a Deltamax core and a resistor, Fig. 1, and this voltage is of such a value that the core does not saturate, the voltage across the resistor e_R, will appear as shown in Fig. 2. From this figure it can be seen

---

Paper 55-546, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 29, 1955; made available for printing May 12, 1955.

D. H. SCHAEFER is with the Naval Research Laboratory, Washington, D. C.

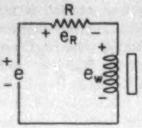


Fig. 1. Circuit comprised of a resistor and a square-loop magnetic core

that when the driving voltage is instantaneously small the core acts like a short circuit. When the driving voltage is instantaneously a large positive value the core limits positive current and when the driving voltage is instantaneously a large negative voltage the core limits negative current. Therefore, before saturation the core assumes three different characteristics: 1. A short circuit. 2. A positive current limiter. 3. A negative current limiter. In terms of the hysteresis loop these three conditions amount to: 1. the point of operation within the boundaries of the hysteresis loop, and 2. and 3. the point of operation on the sides of the loop.

If a measurement of the d-c hysteresis loop of the core used in Fig. 1 is obtained, it will be found that the sides of the loops are practically vertical. It is therefore of interest to determine the characteristics of an idealized core with absolutely vertical aides and eddy-current paths of constant resistance and to see how well data such as shown in Fig. 1 are approximated by such a core. The idealized core will be represented by the complete diagram shown in Fig. 3. The secondary winding of resistance K represents the eddy-current paths in the core itself. These eddy-current paths are an integral part of the core, and in setting up circuit equations the method to be presented will require that these paths be initially assumed present. The eddy-current secondary will therefore also be assumed in all circuit diagrams though only explicitly shown in Fig. 3.

It is desired to obtain an expression for e_w, the voltage across the windings of the idealized core of Fig. 3, in terms of i, the current through the core winding. The current i_s is assumed to represent the eddy currents in the core itself. The half-width of the core hysteresis loop related to the primary circuit is assumed to be I amperes.

If i is less than I, the core is a short circuit and

$$e_w = 0 \tag{1}$$

If i is greater than I, the core is limiting positive current. Under these conditions the total current seen by the core is constrained to the value I. Therefore

$$i - i_s = I \tag{2}$$

also

$$e_w = Ki_s \tag{3}$$

therefore

$$e_w = K(i-I) \tag{4}$$

If i is a negative current less than $-I$, the core is limiting negative current and as the total current is constrained to $-I$

$$i - i_s = -I \tag{5}$$

also

$$e_w = Ki_s \tag{6}$$

therefore

$$e_w = K(i+I) \tag{7}$$

Equation 4 is valid only when $K(i-I)$ is positive. Equation 7 is valid only when $K(i+I)$ is negative. Therefore equations 1, 4, and 7 may be combined in the following fashion

$$e_w = [0\,V K(i-I)]\Lambda K(i+I) \tag{8}$$

where the symbols V and Λ are defined as

$a\,V b$ (read a max b) $=$ max (a, b) $=$ the more positive (or less negative) of a or b

$a \Lambda b$ (read a min b) $=$ min (a, b) $=$ the more negative (or less positive) of a or b

As examples

$2V1=2,\ -2V-1=-1,\ 2\Lambda 1=1,\ -2\Lambda -1=-2.$

With these symbols an algebra somewhat like Boolean algebra can be formed. The basic rules of this mathematics are given in Appendix I. The motivation for this symbolism and algebra was based on the infinite-valued logic of Lukasiewicz and Tarski.[2] The V sign is a generalization of the binary "or," the Λ sign of the binary "and," and the $-$ sign of the binary "not" of classical propositional calculus. Making these translations renders rules 81 to 84, and 86 identical to rules of 2-valued propositional calculus.

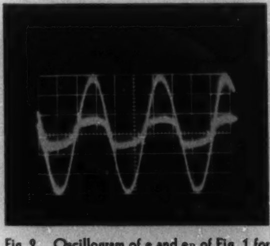


Fig. 2. Oscillogram of e and e_R of Fig. 1 for a Deltamax core and a 60-cycle sinusoidal e not large enough to saturate core

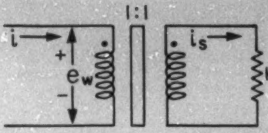


Fig. 3. Equivalent circuit for physical core

To return to the cores, remember that Fig. 3 in its entirety represents a core. The voltage e_R of Fig. 1 as a function of the driving voltage e will be calculated. To do this it is necessary to calculate the voltage when the core is in each of its three states.

First, if the core is a short circuit

$$e = e_R \tag{9}$$

or

$$e_R\langle 0\rangle = e \tag{10}$$

Here $e\langle 0\rangle$ is introduced and defined as the symbol for the voltage across a component when the core has the characteristic of a short circuit.

If the core is limiting positive current

$$e = iR + K(i-I) \tag{11}$$

by Kirchoff's laws and equation 4.

When equation 11 is solved

$$e_R\langle +\rangle = iR = \frac{e+KI}{R+K}R \tag{12}$$

where $e\langle +\rangle$ is by definition the symbol for the voltage across a component, when the core is limiting positive current. Similarly by use of Kirchoff's laws and equation 7 it is found that

$$e_R\langle -\rangle = \frac{e-KI}{R+K}R \tag{13}$$

where $e\langle -\rangle$ is by definition the symbol for the voltage across a component when the core is limiting negative current.

When equations 10, 12, and 13 are combined according to rules to be given in the next section

$$e_R = \left(e\Lambda\frac{e+KI}{R+K}R\right)V\frac{e-KI}{R+K}R \tag{14}$$

or if K is much greater than R

$$e_R = (e\Lambda RI)\,V-RI \tag{15}$$

Fig. 4 shows plots of e and e_R if e is a sinusoid. Fig. 4(A) is the plot that would be obtained from equation 14 if K were a constant larger than but comparable to R, while Fig. 4(B) is the plot that is

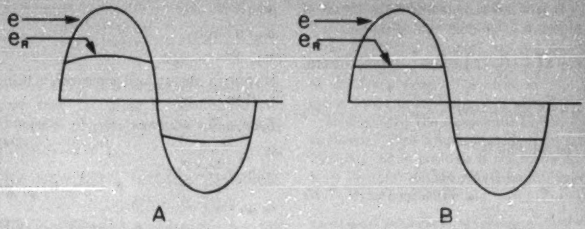


Fig. 4. Plot of e_R

A—From equation 14
B—From equation 15

obtained from equation 15 where K is assumed orders of magnitude greater than R.

When Figs. 2 and 4 are compared it can be seen that qualitatively equation 14 expresses the relationship between e and e_R. However, the process of measuring K results in some surprises.

The previously introduced equation

$$e_w = [0\,V K(i-I)]\Lambda K(i+I) \tag{8}$$

states that if a step of current is passed through a core, there will be: 1. no voltage across the core if the magnitude of this step of current is less than I, and 2. the voltage will be proportional to $(i-I)$ if i is greater than I. A good method of measuring both I and K therefore is to apply step currents to a core.

If square current pulses are actually sent through the windings of a Deltamax core, the results are as follows: when the height of the current pulse is less than I, equation 8 is correct as no voltage will appear across the core. With Deltamax, the value of I is well defined. If the magnitude of the current pulse is made greater than I, however, equation 8 with K assumed a constant will not predict the result. Fig. 5 shows the current through and the voltage across the windings of a Deltamax core that is being driven from negative saturation to positive saturation by a square pulse of current. If K is a constant, equation 8 predicts a voltage that is a square wave. The voltage wave form of Fig. 5 shows that K is not a constant, but rather a quantity that varies as the flux level of the core changes. Pictures similar to Fig. 5 but with current pulses of different heights show that K is also a function of the magnitude of the current pulse. Table I gives various measurements of the quantity K, which, of course, is simply the ratio of the voltage to the quantity $(i-I)$.

The important items to be culled from the experimental results with current pulses are: 1. there is a well-defined I, and 2. equation 8 holds if K is thought of as a variable quantity, i.e., the voltage across the windings of such a core is dependent on the value of $(i-I)$ rather than on the value of di/dt as in an inductance. In fact, Fig. 5 shows the value of the voltage across the core windings when di/dt is zero. This means that a square-loop core is a dissipative element rather than an energy storage element.

If K varies between two values much larger than R, it can be seen that an expression like $KR/K+R$ will very nearly equal R even though K is variable. In Fig. 2 the peak value of e_R is very close to that predicted by equation 15.

For completeness, it should be stated that in a Deltamax core, in addition to the square-loop characteristics treated in the foregoing there is some inductance, both self and mutual, that appears always to be present. At 60 cycles per second, this seems to be of little importance in average magnetic-amplifier circuitry. In the sample tested, which had pie windings on opposite sides of the core, the self-inductance of the core before limiting was on the order of 40 millihenrys, and the mutual inductance was on the order of 5 millihenrys. After saturation, these values are reduced by a factor of 3 or 4.

One-Core Circuits

The rule for finding a voltage e_n in a circuit containing only resistors and one square loop core is: Let e_w be the voltage across the core windings. Determine whether $\partial e_n/\partial e_w$ is positive or negative. In other words, treating the core as a voltage source determines if an increase in this voltage source will increase or decrease e_n, all other voltages being held constant.

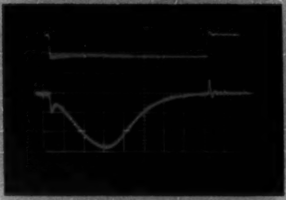


Fig. 5. Voltage wave form showing K not a constant

Top trace—Step pulse of current through primary core windings
Bottom trace—Voltage across open-circuited secondary winding

If $\partial e_n/\partial e_w$ is positive

$$e_n = (e_n\langle 0\rangle V e_n\langle +\rangle)\Lambda e_n\langle -\rangle \tag{16}$$

If $\partial e_n/\partial e_w$ is negative

$$e_n = (e_n\langle 0\rangle \Lambda e_n\langle +\rangle) V e_n\langle -\rangle \tag{17}$$

This rule is proved in Appendix II in a generalized form applicable to any piecewise or approximately piecewise linear element, be the element a rectifier, vacuum tube, transistor, or some other component. The rule will be applied to rectifiers later.

In the case of square-loop cores, it can be shown that the rules and equations 16 and 17 can be given as

$$e_n = [e_n\langle 0\rangle \Lambda(e_n\langle +\rangle V e_n\langle -\rangle)]V\times (e_n\langle -\rangle \Lambda e_n\langle +\rangle) \tag{18}$$

If equation 18 is utilized, it is not necessary to determine the sign of $\partial e_n/\partial e_w$. This latter determination, however, is usually trivial.

As a first example, the voltage across the windings of the core of Fig. 1 will be solved. In this case

$$e_w\langle +\rangle = K(i-I) \tag{19}$$

$$= K\frac{e+KI}{R+K} - KI \tag{20}$$

$$= e\frac{K}{R+K} - \frac{RK}{R+K}I \tag{21}$$

$$= e - RI, \tag{22}$$

if K is much larger than R.

Table I. Measurements of K*

i−I	Maximum Value of K	Initial Value of K
1.8 milliamperes	1,900 ohms	250 ohms
4.2 milliamperes	4,700 ohms	900 ohms

*Deltamax core, 1-mil laminations 500 turns, $I = 2.2$ milliamperes

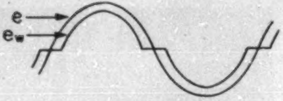


Fig. 6. Plot of e_w from equation 19 for sinusoidal e

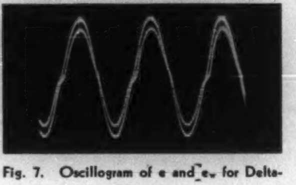


Fig. 7. Oscillogram of e and e_w for Deltamax core

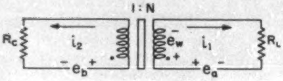


Fig. 8. A 2-loop circuit

Similarly

$$e_w<->=e\frac{K}{R+K}+\frac{RK}{R+K}I \qquad (23)$$

$$=e+RI, \qquad (24)$$

if K is much larger than R.

And $e_w<0>=0$. (25)

Since $e_n=e_w$, $\partial e_n/\partial e_w$ is positive. Therefore equation 16 applies and

$$e_w=[0\vee(e-RI)]\wedge(e+RI) \qquad (26)$$

if K is much larger than R.

It should be noted that $\partial e_R/\partial e_w$ is a negative number, where e_R is the voltage across the resistor of Fig. 1, which accounts for the form of equation 14 in the preceding section.

If $e=\sin \omega t$, e_w of equation 26 appears as shown in Fig. 6. In Fig. 7 e and e_w of a Deltamax core are shown, and the effect of the small residual mutual inductance can be seen as the experimental e_w is the voltage across an open-circuited secondary of the core.

Equation 8 is easily derived for the more general case of a core with two windings, N_p turns on the primary and N_s turns on the secondary. In this instance

$$e_p=\left[0\vee K\left(i_p+\frac{N_s}{N_p}i_s-I\right)\right]\wedge K\left(i_p+\frac{N_s}{N_p}i_s+I\right) \qquad (27)$$

where

e_p = primary voltage
i_p = primary current
i_s = secondary current
I = half-width of loop as measured in the primary circuit

It is of value to consider the circuit of of Fig. 8. To determine $e<+>$

$$e_a=K\left(i_1+\frac{i_2}{N}-I\right)+R_Li_1 \qquad (28)$$

$$e_b=\frac{K}{N^2}(Ni_1+i_2-NI)+R_ci_2 \qquad (29)$$

The solution is

$$e_w=[0\wedge(\alpha e_a+\gamma Ne_b+\theta I)]\times \vee(\alpha e_a+\gamma Ne_b-\theta I) \qquad (30)$$

where

$$\alpha=\frac{N^2R_c}{N^2R_c+R_L}$$

$$\gamma=\frac{R_L}{N^2R_c+R_L}$$

$$\theta=\frac{N^2R_cR_L}{N^2R_c+R_L}$$

if K is much greater than any of the R's.

Solving for the voltage across R_L

$$e_L=[e_a\wedge\gamma(e_a-Ne_b)+\theta I]\vee\gamma(e_a-Ne_b)-\theta I \qquad (31)$$

These equations will be used later.

Two-Core Circuits

When writing equations for circuits containing two cores, the methods given are generalized. A 2-core circuit can operate in nine possible modes since for every state that one core is in, the other can be in any of its three states. These nine modes will be denoted by an $e< >$ with two symbols inside the $< >$, the first referring to the state of core I, the second to the state of core II. For example, $e<0, +>$ is the value of e if core I is in its short-circuited condition and core II is limiting positive current. Further, let the symbol $e<0, \text{any}>$ be the value of e if core I is its short-circuited condition and nothing is specified as to the state of core II. The rules for writing equations for a 2-core circuit are as follows:

1. Find the core in the circuit where $\partial e_n/\partial e_w$ has the same sign, no matter what the state of the other core. Label this core I. In most circuits both cores meet this specification.

2. Determine the sign of $\partial e_n/\partial e_1$ where e_1 is the voltage across the core selected as core I. If this value is positive

$$e_n=(e_n<0, \text{any}>\vee e_n<+, \text{any}>)\wedge e_n<-, \text{any}> \qquad (32)$$

If $\partial e_n/\partial e_1$ is negative

$$e_n=(e_n<0, \text{any}>\wedge e_n<+, \text{any}>)\vee e_n<-, \text{any}> \qquad (33)$$

These equations are completely analogous to equations 16 and 17. In equations 32 and 33 if

$$\frac{\partial e_n<0, \text{any}>}{\partial e_2}$$

is positive, the voltage across core 2 being e_2, then

$$e_n<0, \text{any}>=(e_n<0, 0>\vee e_n<0, +>)\times \wedge e_n<0, -> \qquad (34)$$

while if $\frac{\partial e_n<0, \text{any}>}{\partial e_2}$ is minus

$$e_n<0, \text{any}>=(e_n<0, 0>\wedge e_n<0, +>)\times \vee e_n<0, -> \qquad (35)$$

and similarly for $e<+, \text{any}>$ and $e<-, \text{any}>$.

An illustration should make the foregoing clearer. In Fig. 9(A) it is desired to solve for e_R the voltage across R_1. Here the labeling of the cores can be seen to be proper as $\partial e_R/\partial e_1$ is negative no matter what the state of core II. By equation 33

$$e_R=(e_R<0, \text{any}>\wedge e_R<+, \text{any}>)\vee e_R<-, \text{any}> \qquad (36)$$

the partials $\frac{\partial e_R<0, \text{any}>}{\partial e_2}$, $\frac{\partial e_R<+ \text{any}>}{\partial e_2}$, and $\frac{\partial e_R<-, \text{any}>}{\partial e_2}$

are all positive so that

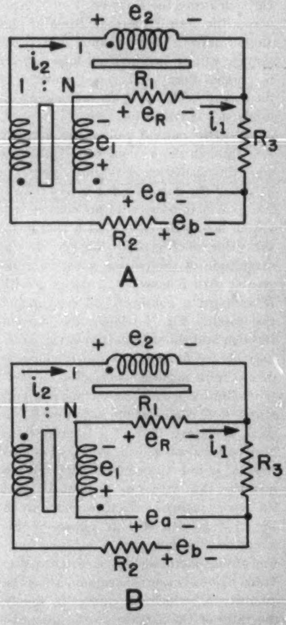


Fig. 9. Circuits containing two cores

$$e_R=[e_R<0,0>Ve_R<0,+>]\Lambda e_R<0,->.\Lambda.[e_R<+,0>Ve_R<+,+>]\Lambda e_R<+,->:V:[e_R<-,0>Ve_R<-,+>]\Lambda e_R<-,-> \quad (37)$$

where a dot notation has been used which can take the place of parenthesis and brackets. In this notation the quantity

$$(aVb)\Lambda c=aVb.\Lambda c \quad (38)$$

the dot replacing the pair of parenthesis. For a more complicated expression

$$[(aVb)\Lambda c]\,Vd=aVb.\Lambda c:Vd \quad (39)$$

the dot pair replacing the pair of brackets. In the remainder of the paper this dot notation will be used, along with parenthesis and brackets.

In Fig. 9(A)

$$e_a=K_1\left(i_1+\frac{i_2}{N}-I_1\right)+(R_1+R_3)i_1+R_3i_2 \quad (40)$$

$$e_b=\frac{K_1}{N^2}(Ni_1+i_2-NI_1)+(R_2+R_3)i_2+R_3i_1+\frac{K_2}{N^2}(i_2-NI_2) \quad (41)$$

To solve for the nine combinations

$$e_R<0,0>=\alpha e_a-\gamma e_b \quad (42)$$

$$e_R<0,+>=\alpha' e_a-\theta NI_2 \quad (43)$$

$$e_R<0,->=\alpha' e_a+\theta NI_2 \quad (44)$$

$$e_R<+,0>=\gamma'(e_a-Ne_b)+\Omega I_1 \quad (45)$$

$$e_R<-0>=\gamma'(e_a-Ne_b)-\Omega I_1 \quad (46)$$

$$e_R<+\ +>=(I_1-I_2)R_1 \quad (47)$$

$$e_R<+\ ->=(I_1+I_2)R_1 \quad (48)$$

$$e_R<-\ +>=(-I_1-I_2)R_1 \quad (49)$$

$$e_R<-\ ->=(-I_1+I_2)R_1 \quad (50)$$

where

$$\alpha=\frac{R_2+R_3}{R_1R_2+R_1R_3+R_2R_3}R_1$$

$$\gamma=\frac{R_3R_1}{R_1R_2+R_1R_3+R_2R_3}$$

$$\alpha'=\frac{R_1}{R_1+R_3}$$

$$\gamma'=\frac{R_1}{R_1+N^2R_2+(N^2-2N+1)R_3}$$

$$\theta=\frac{R_1R_3}{R_1+R_3}$$

$$\Omega=\frac{R_1[N^2R_2+N(N-1)R_3]}{R_1+N^2R_2+(N^2-2N+1)R_3}$$

if K_1 and K_2 are large compared to the R's.

Substituting into equation 37

$$e_R=[(\alpha e_a-\gamma e_b)V(\alpha' e_a-\theta NI_2)]\Lambda\alpha' e_a+\theta NI_2.\Lambda.[\gamma'(e_a-Ne_b)+\Omega I_1V(I_1-I_2)R_1]\Lambda(I_1+I_2)R_1:V:[\gamma'(e_a-Ne_b)-\Omega I_1V(-I_1-I_2)R_1]\Lambda(-I_1+I_2)R_1 \quad (51)$$

This equation was formed assuming cores I and II to be as designated in Fig. 9(A). For this particular circuit, the designations could be reversed as $\partial e_R/\partial e_2$ is always positive, no matter what the state of the lower core. Therefore in Fig. 9 either core can be designated as core I.

In Fig. 9(B) the cores must be labeled as shown. This is necessitated by the fact that $\partial e_R/\partial e_2$ is positive if the lower core is in its short-circuit state, but negative if the lower core is limiting. In Fig. 9 (B)

$$e_R=[e_R<0,0>Ve_R<0,+>]\Lambda e_R<0,->.\Lambda.[e_R<+,0>\Lambda e_R<+,+>]Ve_R<+,->:V:[e_R<-,0>\Lambda e_R<-,+>]Ve_R<-,-> \quad (52)$$

where the nine e's are not entirely identical to equations 42 to 50.

Series-Connected Magnetic Amplifier

As an illustration of the use of the method just described, the equations for the voltage across a core winding of a standard series-connected magnetic amplifier, Fig. 10, will be derived.

If a 1-to-1 turns ratio is assumed

$$e_{ac}=K_1(i_1+i_2-I_1)+R_Li_1+K_2(i_1-i_2-I_2) \quad (53)$$

and

$$e_c=R_ci_2+K_1(i_1+i_2-I_1)+K_2(i_2-i_1+I_2) \quad (54)$$

from which can be obtained

$$e_1<+,+>=\frac{(e_{ac}+K_1I_1+K_2I_2)(R_c+K_1+K_2)-(K_1-K_2)(e_c+K_1I_1-K_2I_2)+(K_1+K_2+R_L)(e_c+K_1I_1-K_2I_2)-(e_{ac}+K_1I_1+K_2I_2)(K_1-K_2)}{(K_1+K_2+R_L)(K_1+K_2+R_c)-(K_1-K_2)^2K_1-K_1I_1} \quad (55)$$

If K_1 is set equal to K_2 and the K's are assumed much larger than the R's, the expression simplifies to

$$e_1<+,+>=\frac{e_{ac}+e_c}{2}-\frac{R_c+R_L}{4}I_1+\frac{R_c-R_L}{4}I_2 \quad (56)$$

and if $I_1=I_2$

$$e_1<+,+>=\frac{e_{ac}+e_c}{2}-\frac{R_L}{2}I \quad (57)$$

Similarly

$$e_1<+,->=\frac{e_{ac}+e_c}{2}-\frac{R_c}{2}I \quad (58)$$

$$e_1<-,+>=\frac{e_{ac}+e_c}{2}+\frac{R_c}{2}I \quad (59)$$

$$e_1<-,->=\frac{e_{ac}+e_c}{2}+\frac{R_L}{2}I \quad (60)$$

also

$$e_1<+,0>=\frac{e_{ac}R_c+e_cR_L}{R_c+R_L}-\frac{R_LR_c}{R_c+R_L}I \quad (61)$$

$$e_1<-,0>=\frac{e_{ac}R_c+e_cR_L}{R_c+R_L}+\frac{R_LR_c}{R_c+R_L}I \quad (62)$$

and $e_1<0,0>=e_1<0,+>=e_1<0,->=0 \quad (63)$

As $\partial e_1/\partial e_1$ is positive, equation 32 is valid. The polarities of $\partial e_1/\partial e_2$ depend upon the relative values of R_c and R_L. If R_c is greater than R_L an increase of e_2 will decrease e_1, or $\partial e_1/\partial e_2$ is negative. On the other hand if R_L is greater than R_c, $\partial e_1/\partial e_2$ is positive.

If R_c is greater than R_L, for example, from equations 32 and 35

$$e_1=0V.\left(\frac{e_{ac}R_c+e_cR_L}{R_c+R_L}-\frac{R_LR_c}{R_L+R_c}I\Lambda\frac{e_{ac}+e_c}{2}-\frac{R_L}{2}I\right)\times V\frac{e_{ac}+e_c}{2}-\frac{R_c}{2}I:\Lambda:\left(\frac{e_{ac}R_c+e_cR_L}{R_c+R_L}+\frac{R_LR_c}{R_L+R_c}I\Lambda\frac{e_{ac}+e_c}{2}+\frac{R_c}{2}I\right)V\frac{e_{ac}+e_c}{2}+\frac{R_L}{2}I \quad (64)$$

so long as both cores are unsaturated. If core II is saturated

$$e_1=\left(0V\frac{e_{ac}R_c+e_cR_L}{R_c+R_L}-\frac{R_LR_c}{R_L+R_c}I\right)\Lambda\times\frac{e_{ac}R_c+e_cR_L}{R_c+R_L}+\frac{R_LR_c}{R_L+R_c}I \quad (65)$$

Circuits Containing Rectifiers

A rectifier can be thought of as an element that has two linear states. In one

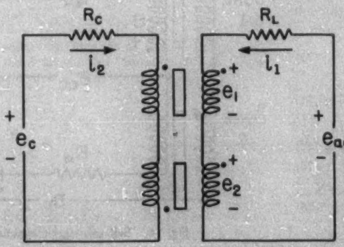


Fig. 10 (left). Series-connected magnetic amplifier

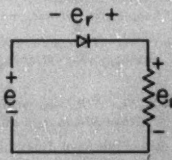


Fig. 11 (below). A rectifier circuit

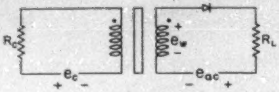


Fig. 12. Basic single-core magnetic amplifier

state it is a short circuit and in the other an open circuit. In a circuit containing only one rectifier, a voltage e_n will be either e_n<open>, the value of e_n calculated for the rectifier considered an open circuit, or e_n<short>, the value calculated for e_n with the rectifier considered a short circuit.

Following the scheme set up for cores, to find a voltage e_n in a circuit containing a rectifier, determine the sign of $\partial e_n/\partial e_r$ where e_r is the rectifier voltage, the plate end of the rectifier being considered positive. In other words, determine whether an increase in forward current through the rectifier (all driving voltages beind held constant) will increase or decrease e_n.

If $\partial e_n/\partial e_r$ is positive

$$e_n = e_n\text{<open>}\, V e_n\text{<short>} \tag{66}$$

If $\partial e_n/\partial e_r$ is negative

$$e_n = e_n\text{<open>}\, \Lambda e_n\text{<short>} \tag{67}$$

A trivial example is Fig. 11. Here

$$e_R\text{<open>} = 0 \tag{68}$$

$$e_R\text{<short>} = e \tag{69}$$

and an increase in forward current through the rectifier will increase e_R. Therefore

$$e_R = 0\, V e \tag{70}$$

The single-core magnetic amplifier, Fig. 12, is a less trivial example. It is of interest to solve for the voltage across the core windings. Here

$$e_w\text{<open>} = (0\, V Ne_c - N^2R_cI)\, \Lambda Ne_c + N^2R_cI \tag{71}$$

from a calculation similar to that for equation 26, and

$$e_w\text{<short>} = (0\, V \alpha e_{ac} + \gamma Ne_c - \theta I)\, \Lambda \alpha e_{ac} + \gamma Ne_c + \theta I \tag{72}$$

from equation 30.

An increase in forward current through the rectifier will decrease e_w. Therefore

$$e_w = e_w\text{<open>}\, \Lambda e_w\text{<short>} \tag{73}$$

$$e_w = (0\, V Ne_c - N^2R_cI)\, \Lambda Ne_c + N^2R_cI : \Lambda : (0\, V \alpha e_{ac} + \gamma Ne_c - \theta I)\, \Lambda \alpha e_{ac} + \gamma Ne_c + \theta I \tag{74}$$

$$= (0\, V Ne_c - N^2R_cI)\, \Lambda (Ne_c + N^2R_cI)\, \Lambda (0\, V \alpha e_{ac} + \gamma Ne_c - \theta I)\, \Lambda (\alpha e_{ac} + \gamma Ne_c + \theta I) \tag{75}$$

$$= 0\, V (Ne_c - N^2R_cI\, \Lambda \alpha e_{ac} + \gamma Ne_c - \theta I)\, .\Lambda\, N^2e_c + N^2R_cI\, \Lambda \alpha e_{ac} + \gamma Ne_c + \theta I \tag{76}$$

by dual of rule 82 of Appendix I.

Circuits containing more than one rectifier are built up by considering the rectifiers one at a time. The expression e<open> and e<short> in a multirectifier circuit will, in general, contain V and Λ signs due to the other rectifiers. A 2-rectifier circuit, the Ramey single-core magnetic amplifier,[3] is examined in Appendix III.

Often circuits containing square-loop cores work in a manner contrary to intuition. The reader is asked to specify in his own mind when the rectifier shown in Fig. 13 will be conducting before examining the following algebraic solution. This circuit is a part of the self-saturating magnetic amplifier with positive feedback that is further examined in Appendix IV.

In Fig. 13 if the rectifier is open

$$e_L\text{<open>} = (\alpha_1 e_{ac} \Lambda - \alpha_2 e_{ac} + R_LI)\, V - \alpha_2 e_{ac} - R_LI \tag{77}$$

where

$$\alpha_1 = \frac{(R_W + R_L)[NR_L - (R_W + R_L)] + N^2R_cR_L}{(R_W + R_L)[(N-1)(R_W + R_L + N^2R_c)]}$$

and

$$\alpha_2 = \frac{R_W}{(N-1)(R_L + R_W) + N^2R_c}$$

In these equations the quantity $R_W + R_L/NR_L$ has been substituted throughout for the feedback factor x.

$$e_L\text{<closed>} = 0$$

Therefore, as forward current through the rectifier increases e_L

$$e_L = [(\alpha_1 e_{ac} \Lambda - \alpha_2 e_{ac} + R_LI)\, V - \alpha_2 e_{ac} - R_LI]\, V 0 \tag{78}$$

By either equation 77 or equation 78 it can be seen that if e_{ac} has the polarity shown and is greater than R_LI/α_2, the rectifier will be conducting. An easy check on equation 77 is to note that if R_c is infinite, the equation degenerates to the form of equation 15. The fact that the rectifier does conduct for large value of e_{ac} means that rectifiers 2 and 3 of the self-saturating magnetic amplifier with feedback shown in Fig. 14 are conducting most of the time when e_{ac} is positive.

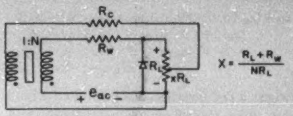


Fig. 13. When will the rectifier be conducting?

A circuit like that in Fig. 14 contains two cores and four rectifiers. It is conceivable that to solve for certain voltages in this circuit it would be necessary to evolve an expression containing 144 elements connected by V and Λ signs as a circuit containing n rectifiers and m cores can operate in 2^n3^m states. Therefore it is desirable to have at least a somewhat quicker method than the complete procedure that has been presented thus far in this paper.

One method of reducing the complexity of the system is to ignore magnetizing current when solving magnetic-amplifier circuits. This is legitimate if all IR terms are small compared to the driving voltages. If in equation 27 I is set equal to zero

$$e_p = \left[0\, V K\left(i_1 + \frac{N_s}{N_p} i_2\right)\right] \Lambda K\left(i_1 + \frac{N_s}{N_p} i_2\right) \tag{79}$$

which simplifies to

$$e_p = K\left(i_1 + \frac{N_2}{N_1} i_2\right) \tag{80}$$

by equation 86.

By making this simplification, all V and Λ signs in the core equations may be discarded. Therefore not more than 2^n equations for an n rectifier circuit need be solved. In Appendix IV the self-

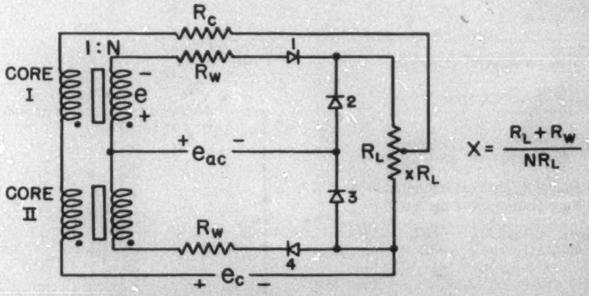


Fig. 14. Self-saturating magnetic amplifier with negative feedback

saturating magnetic amplifier with feedback[4] is analyzed with the use of the simplified equations.

Conclusion

In this paper an attempt is made to show that by using mathematical language containing the ideas of "minimum" and "maximum," it is possible to solve for voltages in a circuit containing square-loop cores and rectifiers in a manner analogous to such solutions for linear circuits.

Through the use of the methods presented here, the operation of unbuilt circuits can be anticipated, equivalences can be concisely defined, demonstrated, or disproved, and physical simplification can often be effected. Most important, the details of operation of a circuit are clearly indicated by the resultant expressions.

The ideas contained herein are generally applicable to all sorts of devices that can be considered piecewise linear. Examples of elements that fit into this category are vacuum tubes, transistors, and nonlinear dielectric materials as well as the nonlinear magnetic materials and resistors considered in this paper.

ol. 74
. 1
955

Appendix I. Basic Relationship Governing V and Λ Signs

All relationships have duals where V's are replaced by Λ's and conversely.

$$(aVb)Vc = aV(bVc) = (aVc)Vb = aVbVc \tag{81}$$

$$(aVb)\Lambda c = (a\Lambda c)V(b\Lambda c) \tag{82}$$

$$aVa = a \tag{83}$$

$$-(aVb) = (-a)\Lambda(-b) \tag{84}$$

$$a+(bVc) = (a+b)V(a+c) \tag{85}$$

$$aV(b\Lambda a) = a \tag{86}$$

$$aV(b\Lambda c) = aVc \tag{87}$$

if c has a value between or less than a and b.

Proof of equation 87:

Given $aVbVc = aVb$

Therefore $(aVbVc)\Lambda(aVc) = (aVb)\Lambda(aVc)$

or $aVc = (aVb)\Lambda(aVc)$

by equation 86

$aVc = aV(b\Lambda c)$ Q.E.D.

by equation 82.

Appendix II. Proof of Basic Circuit Theorem

Given: A piecewise linear element (ple) of k states with a voltage-current characteristic of the following form

$$e_{\text{ple}} = F[f_1(i), f_2(i) \ldots f_k(i)] \tag{88}$$

where F is a function of V and Λ signs only, and the f's are linear functions of i.

To prove: If the element is in a network composed otherwise of resistors, e_n is the voltage across resistor number n and $e_n<k>$ is the voltage across this resistor assuming the piecewise linear element to be in state k, then

$$e_n = F(e_n<1>, e_n<2> \ldots e_n<k>) \tag{89}$$

if $\partial e_n/\partial e_{\text{ple}}$ is positive, or

$$e_n = F^{-1}(e_n<1>, e_n<2> \ldots e_n<k>) \tag{90}$$

if $\partial e_n/\partial e_{\text{ple}}$ is negative, where F^{-1} is the dual of F (by dual is meant that all V's are replaced by Λ's and all Λ's by V's).

Proof: For a piecewise linear circuit, it is possible to solve within each linear region for $i<1>$, $i<2>$, … $i<k>$ by Kirchoff's laws where i is the current flowing through the piecewise linear element. The $i<>$'s will be functions of the driving voltages (the e_i's) in the circuit. Therefore by substitution of the current solutions into equation 88.

$$e = F\{f_1[i<1>(e_i\text{'s})] \ldots f_k[i<k>(e_i\text{'s})]\} \tag{91}$$

$$= F(g<1> \ldots g<k>) \tag{92}$$

where the $g< >$'s are linear functions of the driving voltages.

By use of the compensation theorem the piecewise linear element can be replaced by a voltage generator whose voltage e equals $F(g<1> \ldots g<k>)$. By the superposition theorem it therefore follows that

$$e_n = a_0F(g<1> \ldots g<k>) + \Sigma a_ie_i \tag{93}$$

where the a's are constants and the e_i's are the driving voltages of the system.

Equation 93 may be rewritten

$$e_n = F(a_0g<1> \ldots a_0g<k>) + \Sigma a_ie_i \tag{94}$$

if a_0 is positive. If a_0 is negative

$$e_n = F^{-1}(a_0g<1> \ldots a_0g<k>) + \Sigma a_ie_i \tag{95}$$

as $-(aVb) = (-a)\Lambda(-b)$.

With the equation $a+(bVc) = (a+b)V(a+c)$ in mind, equation 94 can then be written

$$e_n = F(a_0g<1> + \Sigma a_ie_i \ldots a_0g<k> + \Sigma a_ie_i) \tag{96}$$

$$= F(e_n<1> \ldots e_n<k>) \tag{97}$$

and similarly equation 95 becomes

$$e_n = F^{-1}(e_n<1> \ldots e_n<k>) \tag{98}$$

where the sign of a_0 (i.e., $\partial e_n/\partial e_{\text{ple}}$) determines whether equation 97 or 98 is the desired form.

Appendix III. Ramey Single-Core Amplifier

A 2-rectifier circuit is treated as two single rectifier circuits. In Fig. 15 it is desired to solve for the voltage across the core windings. If e^0 is defined as the voltage across the core winding when both rectifiers are open circuits, e^1 as the voltage across the core winding when rectifier 1 only is a short circuit, e^2 is the voltage when rectifier number 2 only is a short circuit, and e^{12} the voltage when rectifiers 1 and 2 are both short circuits

$$e^0 = 0 \tag{99}$$

$$e^1 = [0V(e_a - R_LI)]\Lambda(e_a + R_LI) \tag{100}$$

$$e^2 = [0V(Ne_b - N^2R_cI)]\Lambda(Ne_b + N^2R_cI) \tag{101}$$

$$e^{12} = [0V(\alpha e_a + \gamma Ne_b - \theta I)]\Lambda(\alpha e_a + \gamma Ne_b + \theta I) \tag{102}$$

where α, γ, and θ are defined in equation 30.

It is possible to use rectifier no. 1 as the base rectifier as increases in forward current through this rectifier will decrease the core voltage no matter what the state of the left-hand rectifier. Therefore

$$e = e<\text{no. 1 open}>\Lambda e<\text{no. 1 short-circuited}> \tag{103}$$

The expression for $e<\text{no. 1 open}>$ is arrived at exactly as in the case of a single rectifier circuit. Therefore

$$e<\text{no. 1 open}> = e^0Ve^2 \tag{104}$$

$$\begin{aligned} &= 0V.[0V(Ne_b - N^2R_cI)] \times \Lambda(Ne_b + N^2R_cI) \\ &= 0V0V(Ne_b - N^2R_cI).\Lambda.0V(Ne_b + N^2R_cI) \end{aligned} \tag{105}$$

by the dual of rule 82

$$= 0V(Ne_b - N^2R_cI).\Lambda.0V(Ne_b + N^2R_cI) \tag{106}$$

by rule 83

$$= 0V(Ne_b - N^2R_cI) \tag{107}$$

as N, R_c, and I are positive numbers.

Similarly

$$e<\text{no. 1 short-circuited}> = e^1Ve^{12} \tag{108}$$

$$e<\text{no. 1 short-circuited}> = [0V(e_a - R_LI)] \times \Lambda(e_a + R_LI).V.[0V(\alpha e_a + \gamma Ne_b - \theta I)]\Lambda(\alpha e_a + \gamma Ne_b + \theta I) \tag{109}$$

Substituting into equation 103

$$e = 0V(Ne_b - N^2R_cI).\Lambda.[0V(e_a - R_LI)] \times \Lambda(e_a + R_LI).V.[0V(\alpha_a + \gamma Ne_b - \theta I)]\Lambda(\alpha e_a + \gamma Ne_b + \theta I) \tag{110}$$

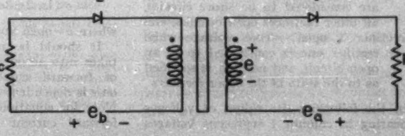


Fig. 15. Ramey single-core magnetic amplifier

$$e=0V(Ne_b-N^2R_cI).\Lambda.0V(e_a-R_LI)\times V(\alpha e_a+\gamma Ne_b-\theta I).\Lambda.0V(e_a-R_LI)V(\alpha e_a+\gamma Ne_b+\theta I).\Lambda.0V\times(e_a+R_LI)V(\alpha e_a+\gamma Ne_b-\theta I).\Lambda.(e_a+R_LI)V(\alpha e_a+\gamma Ne_b+\theta I) \tag{111}$$

by dual of rule 82 of Appendix I

$$=0V(Ne_b-N^2R_cI).\Lambda.0V(e_a-R_LI)\times V(\alpha e_a+\gamma Ne_b-\theta I).\Lambda.(e_a+R_LI)\times\Lambda(\alpha e_a+\gamma Ne_b+\theta I) \tag{112}$$

by inspection as R_L, R_c, and I are positive numbers

$$=0V.(Ne_b-N^2R_cI)\Lambda[(e_a-R_LI)V(\alpha e_a+\gamma Ne_b-\theta I)]:\Lambda.(e_a+R_LI)V(\alpha e_a+\gamma Ne_b+\theta I) \tag{113}$$

At this point it is necessary to note that $\alpha e_a+\gamma Ne_b-\theta I$ always has a value between or greater than e_a-R_LI and $Ne_b-N^2R_cI$. This is dictated by the circuitry.

It is also necessary to note that if c has a value between or greater than a and b, then

$$a\Lambda(bVc)=a\Lambda c \tag{114}$$

by the dual of equation 87.

Therefore equation 113 becomes

$$e=0V[(Ne_b-N^2R_cI)\Lambda(\alpha e_a+\gamma Ne_b-\theta I)]\times.\Lambda.(e_a+R_LI)V(\alpha e_a+\gamma Ne_b+\theta I) \tag{115}$$

In the usual design of a Ramey amplifier $e_a=e_{ac}$ and $e_b=(e_{ac}/N)-e_c$ where e_c is always positive. Substituting into equation 115

$$e=0V[(e_{ac}-Ne_c-N^2R_cI)\Lambda(e_{ac}-\gamma Ne_c-\theta I)].\Lambda.(e_{ac}+R_LI)V(e_{ac}-\gamma Ne_c+\theta I) \tag{116}$$

$$=[0V(e_{ac}-Ne_c-N^2R_cI)]\Lambda(e_{ac}+R_LI) \tag{117}$$

as e_c, R_c, R_L, and I are positive quantities, γ is positive and less than 1, and θ is positive and less than R_L and N^2R_c.

From equation 117 it can be seen, for instance, that if e_{ac} is less than $Ne_c+N^2R_cI$, the core will receive no reset.

Appendix IV. Analysis of Self-Saturating Magnetic Amplifier with Negative Feedback, Neglecting Effects of Magnetizing Currents

Nomenclature

e^0=core voltage when all rectifiers are considered to be open circuits
e^1=core voltage when rectifier 1 is considered to be a short circuit; all other rectifiers open circuits
e^{12}=core voltage when rectifiers 1 and 2 are considered to be short circuits, all other rectifiers open circuits, etc.
e<rectifier 1 open>=core voltage when rectifier one is considered to be an open circuit and nothing is specified as to the state of the other rectifiers

In the following the values of voltages appearing in column 1 represent voltages across core I of Fig. 14 calculated assuming core II unsaturated, and those appearing in column 2 represent voltages across core I calculated assuming core II saturated. In these calculations K_1 and K_2 were assumed much larger than the resistors and the magnetizing currents I_1 and I_2 were assumed zero.

		Column 1	Column 2
$e^0=e^1=e^2=$ $e^3=e^4=e^{23}=$		f_1Ne_c	Ne_c
e^{12}	$=$	e_{ac}	$\alpha e_{ac}+\beta Ne_c$
$e^{13}=e^{134}$	$=$	e_{ac}	e_{ac}
e^{14}	$=$	$f_1\gamma Ne_c$	γNe_c
e^{24}	$=$	$e_{ac}+Ne_c$	$e_{ac}+Ne_c$
$e^{34}=e^{234}$	$=$	$e_{ac}+Ne_c$	Ne_c
e^{123}	$=$	e_{ac}	$\mu e_{ac}+\phi Ne_c$
e^{124}	$=$	$e_{ac}+\gamma Ne_c$	$e_{ac}+\gamma Ne$
e^{1234}	$=$	$e_{ac}+\theta Ne_c$	$\mu e_{ac}+\phi Ne_c$

where

$$\alpha=\frac{NR_L+NR_w+N^2R_c}{NR_L+(N+1)R_w+N^2R_c}$$

$$\beta=\frac{R_w}{NR_L+(N+1)R_w+N^2R_c}$$

$$\gamma=\frac{R_w}{N^2R_c+(N-1)R_L+NR_w}$$

$$\mu=\frac{N^2R_c+(N-1)R_L+(N-2)R_w-\frac{R_w^2}{R_L}}{N^2R_c+(N-1)R_L+(N-1)R_w-\frac{R_w^2}{R_L}}$$

$$\theta=\frac{R_w}{N^2R_c+(N-1)R_L+NR_w-\frac{R_w^2}{R_L}}$$

$$\phi=\frac{R_w}{N^2R_c+(N-1)R_L+(N-1)R_w-\frac{R_w^2}{R_L}}$$

and $f_1=\dfrac{K_1}{K_1+K_2}$

The method of writing equations for a four rectifier circuit is a direct extension of the method for one and two rectifier circuits. In this example the effect of forward current through rectifiers 3 and 4 on the core is dependent on the state of the other rectifier. These rectifiers, therefore, cannot be used as base rectifiers. In this circuit

$$e=e\text{<rectifier no. 1 open>}Ve\text{<rectifier no. 1 short-circuited>} \tag{118}$$

$$=(e\text{<no. 1 open, no. 2 open>}\Lambda e\text{<no. 1 open, no. 2 short-circuited>})V\times(e\text{<no. 1 short-circuited, no. 2 open>}\Lambda e\text{<no. 1 short-circuited, no. 2 short-circuited>}) \tag{119}$$

$$=(e\text{<1o, 2o, 3o>}Ve\text{<1o, 2o, 3s>}.\Lambda.e\text{<1o, 2s, 3o>}\Lambda e\text{<1o, 2s, 3s>})V\times(e\text{<1s, 2o, 3o>}Ve\text{<1s, 2o, 3s>}.\Lambda.e\text{<1s, 2s, 3o>}\Lambda e\text{<1s, 2s, 3s>}) \tag{120}$$

where o=open and s=short.

It should be noted that equation 120 takes into account the fact that the effect of forward current through rectifier 3 on e is dependent on the state of rectifier 2. Note for equation 121 that an increase in forward current through rectifier 4 will decrease e unless both rectifier 2 and 3 are open.

$$e=(e\text{<1o, 2o, 3o, 4o>}?e\text{<1o, 2o, 3o, 4s>}.V.e\text{<1o, 2o, 3s, 4o>}\Lambda e\text{<1o, 2o, 3s, 4s>}:\Lambda:e\text{<1o, 2s, 3o, 4o>}\Lambda e\text{<1o, 2s, 3o, 4s>}.\Lambda.e\text{<1o, 2s, 3s, 4o>}\Lambda e\text{<1o, 2s, 3s, 4s>})V(e\text{<1s, 2o, 3o, 4o>}Ve\text{<1s, 2o, 3o, 4s>}.V.e\text{<1s, 2o, 3s, 4o>}\Lambda e\text{<1s, 2o, 3s, 4s>}:\Lambda:e\text{<1s, 2s, 3o, 4o>}\Lambda e\text{<1s, 2s, 3o, 4s>}.\Lambda.e\text{<1s, 2s, 3s, 4o>}\Lambda e\text{<1s, 2s, 3s, 4s>}) \tag{121}$$

$$=(e^0?e^4.V.e^3\Lambda e^{34}:\Lambda:e^2\Lambda e^{24}\Lambda e^{23}\Lambda e^{234})V\times(e^1Ve^{14}V.e^{13}\Lambda e^{134}:\Lambda:e^{12}\Lambda e^{124}\Lambda e^{123}\times\Lambda e^{1234}) \tag{122}$$

Noting equalities of the various e's, equation 122 reduces to

$$e=[e^0.V.e^0\Lambda e^{34}:\Lambda:e^0\Lambda e^{24}\Lambda e^{34}]V[e^0Ve^{14}Ve^{13}\times.\Lambda.e^{12}\Lambda e^{124}\Lambda e^{123}\Lambda e^{1234}] \tag{123}$$

$$=[e^0\Lambda e^{24}\Lambda e^{34}]V[e^0Ve^{14}Ve^{13}.\Lambda e^{12}\Lambda e^{124}\Lambda e^{123}\times\Lambda e^{1234}] \tag{124}$$

by rule 86 of Appendix I.

Before saturation of core II, the following is true.

$$e=[f_1Ne_c\Lambda e_{ac}+Ne_c]V[f_1Ne_cVf_1\gamma Ne_cVe_{ac}\times.\Lambda e_{ac}\Lambda e_{ac}+\gamma Ne_c\Lambda e_{ac}+\theta Ne_c] \tag{125}$$

$$=[f_1Ne_c\Lambda e_{ac}+Ne_c]V[e_{ac}\Lambda e_{ac}+\gamma Ne_c\Lambda e_{ac}+\theta Ne_c] \tag{126}$$

by rule 86 of Appendix I

$$=[f_1Ne_c\Lambda e_{ac}+Ne_c]V[e_{ac}\Lambda e_{ac}+\theta Ne_c] \tag{127}$$

as $\theta>\gamma>0$.

Equation 127 is the voltage across the upper core of Fig. 14 before saturation of the lower core. In this expression e_c and e_{ac} can be either positive or negative. If e_c is restricted to positive values only, equation 127 becomes

$$e=[f_1Ne_c\Lambda e_{ac}+Ne_c]Ve_{ac} \tag{128}$$

which is the equation for the usual mode of operation of the circuit.

If e_c is restricted to negative values only, equation 127 becomes

$$e=[f_1Ne_c\Lambda e_{ac}+Ne_c]Ve_{ac}+\theta Ne_c \tag{129}$$

$$e=[f_1Ne_cVe_{ac}+\theta Ne_c]\Lambda[e_{ac}+Ne_cVe_{ac}+\theta Ne_c] \tag{130}$$

by rule 82 of Appendix I

$$=[f_1Ne_cVe_{ac}+\theta Ne_c]\Lambda[e_{ac}+\theta Ne_c] \tag{131}$$

as θ is less than 1

$$e=e_{ac}+\theta Ne_c \tag{132}$$

by rule 86 of Appendix I.

After saturation of core II, equation 124 becomes

$$e=[Ne_c\Lambda e_{ac}+Ne_c]V[Ne_cV\gamma Ne_cVe_{ac}\times.\Lambda\alpha e_{ac}+\beta Ne_c\Lambda e_{ac}+\gamma Ne_c\Lambda\mu e_{ac}+\phi Ne_c] \tag{133}$$

Core II usually saturates only when e_{ac} is negative and e_c is positive. Under these special circumstances, equation 133 becomes

Vol.
Pt. 1
1955

$$e=e_{ac}+Ne_cV[Ne_c\Lambda\alpha e_{ac}+\beta Ne_ce_{ac}+\gamma Ne_c] \tag{134}$$

as $\alpha<1$, $\gamma<1$, $\mu<1$, and $\phi>\gamma\beta$

$$=[e_{ac}+Ne_cVNe_c]\Lambda[e_{ac}+Ne_cV\alpha e_{ac}+\beta Ne_c\Lambda e_{ac}+\gamma Ne_c] \tag{135}$$

by rule 82 of Appendix I

$$=[e_{ac}+Ne_cVNe_c]\Lambda[e_{ac}+Ne_cV\alpha e_{ac}+\beta Ne_c.\Lambda.e_{ac}+Ne_cVe_{ac}+\gamma Ne_c] \tag{136}$$

by dual of equation 82

$$=[e_{ac}+Ne_c\,V\,Ne_c]\Lambda[e_{ac}+Ne_c\,V\alpha\,e_{ac}+\beta Ne_c.\Lambda.e_{ac}+Ne_c] \tag{137}$$

as e_c is positive and $\gamma<1$

$$=[e_{ac}+Ne_c\,\Lambda\,Ne_c]\Lambda e_{ac}+Ne_c \tag{138}$$

by equation 86

$$=e_{ac}+Ne_c \tag{139}$$

by equation 86.

References

1. A RECTIFIER ALGEBRA, David H. Schaefer. *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 679–82.

2. SYMBOLIC LOGIC (book), C. I. Lewis, C. H. Langford. Appleton-Century Company, New York, N. Y., 1932, chap. VII (the 3 and n valued form of this logic).

3. ON THE MECHANICS OF MAGNETIC AMPLIFIER OPERATION. Robert A. Ramey. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

4. FAST RESPONSE WITH MAGNETIC AMPLIFIERS, D. G. Scorgie. *Ibid.*, vol. 71, pt I, 1953, p. 741.

Aluminum Conductor Steel-Reinforced (ACSR) for Railroad Signal and Communications Circuits

H. W. ADAMS
MEMBER AIEE

ol. 74
. 1
955

INTEREST in ACSR for railroad signal and communications circuits was stimulated by the copper shortage which occurred during the Korean War and has continued since that time. Because such circuits may be used for communications frequencies up to 150 kc per second as well as for 60-cycle power and direct current, lack of information regarding the communications characteristics of ACSR has retarded its use. In addition, there has been need for accessories and connectors considered suitable and for information regarding mechanical installation and performance. Progress in overcoming these obstacles has now been made to the extent that a test section of line has been constructed by the St. Louis and San Francisco Railway Company.

Communications Characteristics

Only small sizes of ACSR are likely to be used for communication circuits, and investigations thus far have centered upon standard American Wire Gauge (AWG) sizes numbers 4, 6, and 8, all with six aluminum strands over one steel strand. Following methods presented by Green[1] and using measured values of resistance and inductance reported in a 1953 AIEE unpublished paper by M. P. Pyne, values of attenuation were computed for frequencies up to 150 kc per second. The attenuation values computed for dry weather conditions are plotted in Fig. 1 and for wet weather in Fig. 2. For ready comparison, these figures also show attenuation values for commonly used sizes of copper and copperweld conductors.[2] It is interesting to note that the shapes of curves for ACSR are quite similar to those for solid copper wires.

In connection with voice frequency transmission the characteristic impedances of the three sizes of ACSR at 1,000 cycles were calculated. The characteristic impedances, calculated by using values of inductance extrapolated from data presented by M. P. Pyne are 575 ohms for AWG size no. 8, 545 ohms for AWG size no. 6, and 515 ohms for AWG size no. 4.

All of the calculated communications characteristics indicated the feasibility of ACSR and because the attenuation curve of AWG no. 6-6/1 follows the curve of 0.104-inch-diameter copper so closely, the no. 6 size was selected for a test section of line. The section of line chosen for the test is about 8 miles long in the vicinity of Chaffee, Miss., and is part of a completely new section so that it could be tested without interference by tap-offs. Although calculations assumed bare conductor, it is standard practice with the St. Louis and San Francisco Railway to use covered conductors for these circuits. Neoprene-covered ACSR was actually used, therefore, although the effect, if any, that the covering might have on the communications characteristics was not known at the time.

Accessories and Connectors

Because the voltages of communications circuits are quite small and insufficient to bridge any high-resistance connection, it was felt that satisfactory service could be provided only by compression-type splices and connectors. Further, connections to the line at terminals and at taps would normally have to be made to copper wire sizes AWG no. 9 and no. 14, and connectors would have to be suitable for the aluminum-to-copper connections. So far as could be determined, there was not available any standard group of fittings that would fully meet all of the requirements.[3] The most promising line of accessories appeared to be the Crimpit line.[4] This line also had the advantage of ease of installation by means of a small bolt-cutter type of compressor which could be easily handled on the pole by the linemen.

The Burndy Company co-operated in the project by furnishing modifications of some of their standard items to accommodate the sizes of ACSR and copper conductors involved. A special tension splice was designed and furnished for the no. 6-6/1 ACSR which, with only an aluminum sleeve, develops approximately 90 per cent of the breaking strength of the conductor. For simplicity of installation, all tap and terminal connectors were designed to receive no. 8 copper wire, and were actually installed with short

---

Paper 55-559, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 31, 1955; made available for printing May 13, 1955.

H. W. ADAMS is with the Reynolds Metals Company, Louisville, Ky.

The author wishes to acknowledge the co-operation and contributions of R. W. Troth, R. E. Testerman, and R. M. Laurenson, all of the St. Louis-San Francisco Railway Company. Without their active interest and effort this study would not have been possible.

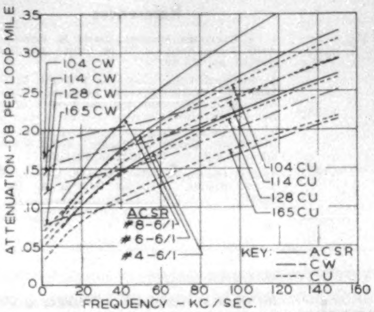


Fig. 1 (left). Dry weather attenuation of ACSR, copper and copperweld conductors; 68 degrees Fahrenheit, 53 pairs of conductor size insulators per mile

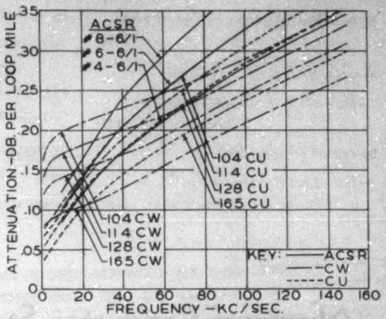


Fig. 2. Wet weather attenuation of ACSR, copper and copperweld conductors; 68 degrees Fahrenheit, 53 pairs of conductor size insulators per mile

no. 8 copper pigtails for future connection of taps and terminals. Tap connectors were of the Eltap type, and butt connectors of the Linkit type. Dead ends were of the loop type, with the conductor being looped around a clevis spool dead-end bracket and secured with Crimpit type compression dead-end fittings. All splices and connectors were installed, using the conventional practice of abrading the surface of the aluminum conductor underneath an overlaying layer of joint compound prior to compression of the fitting. Pentrox *A* joint compound having suspended metallic particles was used.

For attachment to insulators conventional American Telephone and Telegraph Company type of ties were used with neoprene-covered no. 6 solid annealed aluminum tie wire. Since the ACSR was covered with a neoprene jacket, it was not necessary to provide additional protection against abrasion at insulator supports except at dead ends. Aluminum armor tape was used only at dead ends but, had the ACSR been bare, aluminum armor tape or preformed line guards would have been used at all supports.

Sags and Tensions

It was considered desirable, both to simplify construction practices and to provide approximately the same sags as in other types of conductors on the same poles and crossarms, to string the ACSR in to the same sags as are customarily used with other types of conductors. A study was made, therefore, to determine the sags and tensions that would develop in the different sizes of ACSR when so strung in the heavy, medium, and light loading district. A span of 175 feet was selected for computations because ice and wind loadings would have more effect on it than on shorter spans. The ACSR conductors were assumed to be strung initially at 60 degrees Fahrenheit to sags of 21 inches in the heavy and medium loading districts, and of 16 inches in the light loading district. These are the sags set forth in railroad specifications for other types of conductors.[5] Table I shows the sags and tensions that would be developed in a 175-foot span under conditions of full weather loading after being strung to these initial sags, along with the breaking strengths. In parenthesis after the tension in each case is the percentage of the rated breaking strengths that the tension represents. The advantage of the high tensile strength of ACSR is apparent in Table I. The percentage of rated strength developed in each case is quite conservative, except in the case of AWG no. 8 in the heavy loading district, and represents a high degree of safety against breakage during stormy weather. The sags and tensions of commonly used sizes of copper and copperweld conductors similarly computed are as shown in Table II.

In addition to the loaded sags and tensions, the differences in stringing sags and tensions of ACSR from the sags in the railroad specifications through a range of temperatures from 0 to 100 degrees Fahrenheit is of interest. An indication of these differences may be had from Table III which shows the initial sags and tensions in a 175-foot span at various temperatures assuming the ACSR to be strung at 60 degrees Fahrenheit to the specified sag of 21 inches in the heavy loading district. It will be seen that the sags are quite close, as are the tensions for the sizes of conductors shown in the table. These values would be somewhat different if the ACSR were assumed to be strung at some other temperature because of the differences in coefficients of thermal expansion.

The economic advantage of using

Table I. Sags and Tensions in a 175-Foot Span

	Loaded Sags, Inches			Loaded Tensions, Pounds			Breaking Strength, Pounds
ACSR Size	Heavy	Medium	Light	Heavy	Medium	Light	
No. 8—6/1	82	61	37	505 (67%)	369 (49%)	210 (28%)	755
No. 6—6/1	66	51	32	652 (56%)	464 (40%)	290 (25%)	1170
No. 4—6/1	54	42	26	835 (46%)	603 (33%)	379 (21%)	1830

Table II. Sags and Tensions of Copper and Copperweld Conductors

	Loaded Sags, Inches			Loaded Tensions, Pounds			Breaking Strength, Pounds
Conductor Size	Heavy	Medium	Light	Heavy	Medium	Light	
0.104 Copper	75	58	36	520 (95%)	344 (63%)	170 (31%)	550.4
0.114 Copper	70	54	34	565 (86%)	375 (57%)	193 (29%)	661.2
0.128 Copper	64	50	32	626 (76%)	420 (51%)	226 (27%)	826.0
0.165 Copper	53	41	30	794 (60%)	551 (42%)	306 (23%)	1326
0.104 Copperweld	66	52	34	584 (50%)	382 (33%)	182 (16%)	1170
0.114 Copperweld	62	56	32	634 (46%)	524 (38%)	206 (15%)	1368
0.128 Copperweld	57	45	30	697 (43%)	461 (28%)	237 (14%)	1647
0.165 Copperweld	48	38	27	877 (35%)	579 (23%)	337 (13%)	2523

Table III. Sags and Tensions in a 175-Foot Span at Various Temperatures

Temperature, Degrees Fahrenheit	Initial Sag, Inches		Initial Tensions, Pounds	
	ACSR	Railroad Specifications	No. 6–6/1 ACSR	0.114-Inch Copper, Railroad Specifications
0	9	11.5	183	160
20	12.7	14.2	138	135.3
40	16.7	17.3	103	110.3
60	21	21	79	85
80	26.3	28	65	67
100	31.5	35	51	50

ACSR will have to be determined in each individual case. It will depend to a great extent on sizes of different types of conductors being considered and upon prevailing market prices. Table IV gives some idea of the comparative weights per loop mile of various sizes and types of conductors.

Line Construction

Actual installation of the test section of line was accomplished without encountering any unexpected problems. Linemen were able to handle the conductor in their accustomed way and had no serious difficulty with the accessories and connectors.

Attenuation Tests

The attenuation measurements made on this line were intended not to provide absolute data but to indicate whether the performance of the ACSR would be reasonably in accord with the computations. The first measurements were made in December 1953, but showed extremely erratic loss-frequency characteristics. A physical check revealed that construction was incomplete with conductors not being tied to all insulator supports, and there were some line transposition irregularities.

After correction of the irregularities attenuation measurements were again made in February 1955. The developed data are shown in Fig. 3. It will be noted that the measured values are parallel to and slightly below the calculated values up to about 95 kc. Above this frequency the measured attenuation increased rapidly, peaked at 127 kc, and then decreased to approximately the calculated curve at 138 kc. Since the line was transposed to a 3-kc pattern, the attenuation peak between 95 and 138 kc was probably caused by inadequate transpositions which permitted energy to be absorbed in other wires of the line.

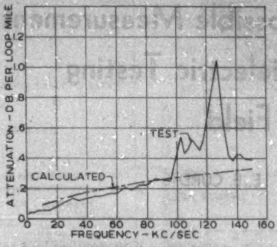


Fig. 3. Measured attenuation of no. 6 ACSR after correction of construction irregularities

Factors contributing to the measured attenuation being lower than computed values at frequencies below 95 kc were the line pin spacing and the test temperature. Computations were for a pin spacing of 8 inches and a temperature of 68 degrees Fahrenheit, whereas the actual pin spacing was 12 inches, and the temperature at the time of the test was about 40 degrees Fahrenheit. Any effect of the neoprene covering was not apparent.

Conclusions

All of the work thus far accomplished indicates that ACSR is technically suitable for railroad signal and communications circuits and gives good promise of advantages in economy and in service continuity. Because the communications characteristics of different types of conductors vary at different rates with changes in frequency, ACSR cannot be compared with copper types of conductors on an equal-conductivity basis as is commonly done for d-c or low-frequency power circuits. A particular size of ACSR, however, may combine to a considerable degree the communications characteristics of a solid copper conductor with the high mechanical strength of a copper-covered steel conductor. Experience with the use of ACSR will no doubt result in expansion and refinement of data relating to its communications characteristics and to installation practices.

Table IV. Weights per Loop Mile of Various Sizes and Types of Conductors

Conductor Size and Type	Weight per Loop Mile, Pounds
No. 8—6/1 ACSR	239.7
No. 6—6/1 ACSR	381.2
No. 4—6/1 ACSR	606.2
0.104 Copper	345.8
0.114 Copper	418.6
0.128 Copper	527.8
0.165 Copper	870.2
0.104 Copperweld	317
0.114 Copperweld	383.6
0.128 Copperweld	480.2
0.165 Copperweld	797.8

References

1. THE TRANSMISSION CHARACTERISTICS OF OPEN WIRE TELEPHONE LINES, E. I. Green. *Bell System Technical Journal*, New York, N. Y., Oct. 1930.

2. ADVANCE NOTICES. Association of American Railroads, New York, N. Y., Communications Section, Oct. 1952.

3. A GUIDE TO THE SELECTION OF ELECTRIC CONDUCTOR ACCESSORIES FOR ALUMINUM WIRES AND CABLES USED IN OVERHEAD TRANSMISSION AND DISTRIBUTION LINES (booklet). Reynolds Metals Company, Louisville, Ky.

4. BULLETIN CR. Burndy Engineering Company, Norwalk, Conn.

5. SPECIFICATIONS FOR THE CONSTRUCTION OF RAILROAD COMMUNICATION POLE LINES, MANUAL SECTION J. Association of American Railroads, New York, N. Y.

The Recognition of Possible Measurement Errors in D-C Dielectric Testing in the Field

C. W. ROSS
ASSOCIATE MEMBER AIEE

E. B. CURDTS
MEMBER AIEE

IN RECENT years, a definite trend appears to have developed, toward d-c testing of insulation. The reasons or this trend are as follows:

1. The d-c test set required is much more portable than an equivalent a-c test set and power requirements are much less.
2. High-voltage d-c tests are less destructive than equivalent a-c tests, as evidenced by the much flatter voltage-time endurance curve in the case of direct current.
3. From the d-c tests important information may be obtained which is related to the condition of the insulation. Periodic tests make it possible to follow the insulation deterioration and thereby aid in determining the need of reconditioning steps and estimating life expectancy.

Discussion

The direct voltage and current relationship as expressed in Ohm's law for electrical conductors is not the same in the case of dielectrics because time enters into this relationship as a parameter. Theoretically, it is only at infinite time that the proportionality constant R, which relates voltage and current in Ohm's law, holds for a dielectric. To express this relationship in the simple terms of Ohm's law it becomes necessary to define R as an apparent resistance at some particular time. It is this time dependency which introduces complications in measurements when an attempt is made to evaluate the condition of a dielectric by subjecting it to d-c electrification.

That valid interpretations may be made from d-c tests regarding the condition of insulation, it is necessary that the theory of the tests be well understood and that certain precautions be followed in making measurements. The time required to make many of the important d-c tests often becomes prohibitively long for many routine maintenance tests because of the long absorption time for many insulations. It is in the attempt to shorten the testing time that many of the pitfalls in d-c measurements occur. It is not difficult to show that many field practices lead to erroneous results. The errors are not obvious when the measurements are being made, but they can cause serious mistakes in interpretation. The purpose of the following is to point out the manner in which certain of these errors are made and how some of them may be recognized from the test data.

Basic Considerations and Test Methods

The interpretation of d-c test data is generally based on one or more of the components of the specimen's charge or discharge current shown in Fig. 1. The geometric charging current component is determined by the electrical impedance R_s of the power source, and the geometric capacitance C of the specimen, generally measured at 60 or 1,000 cycles per second. This component has no direct bearing on the condition of the insulation and should be allowed to reduce to a negligibly small value before significant measurements are made.

The absorption current component is generally given by

$$i_a = VCDt^{-n}$$

D is a constant which is characteristic of the insulation and generally increases because of moisture and the aging of the dielectric.[1–3] The storage coefficient n is usually between 0.5 and 1.0 and may vary somewhat with aging.[3]

The final leakage current is also an indication of moisture, contamination and aging. In some cases a plot of leakage current or resistance versus voltage tends to predict the actual breakdown voltage.[4–6] A pronounced increase in current with time at a given voltage may also indicate an impending failure.

The irreversible absorption current is not measured directly and has not been used in the evaluation of solid insulation, except possibly in the evaluation of the condition of insulating oils. In many practical insulation systems the time required for the absorption current to reach a negligible value may be 1 hour or more. To reduce the testing time, the apparent resistance, or the resistance, may be measured after some specified electrification time to give an indication of the condition of the insulation.

Polarization index, the ratio of the 10-minute apparent resistance to the 1-minute value, is a dimensionless quantity often used in insulation evaluation. Since the leakage current increases at a faster rate with moisture and contamination than does the absorption current, the polarization index is lowered for insulation in the poorer condition. Polarization index is affected very little by temperature.

The 2-voltage resistance test (Evershed or electro-endosmosis effect) is often used to determine the presence of moisture.[7] The principle of this test is the decrease of resistance with increasing voltage. The 2-polarity test is also used to detect moisture.[7] The resistance measured is lower in this test when the center or high-voltage electrode is negative.

Discharge Curves

In some cases it is desirable to know the characteristics of the absorption current component only. The magnitude and rate of decay of this current are related to the moisture content and aging of a given insulation. Sometimes the total charge found by integrating the discharge absorption current from a finite time to infinity, "specific absorption," is used in evaluating insulation.[4] This test may be preferred, especially when the final conduction current is appreciable, or when measuring instrument pointer excursions due to line voltage transients are particularly troublesome during charge.

However, in this type of test it is often not realized that the behavior of the discharge current is greatly affected by the charging time. This may be shown by considering the absorption current only and by use of the superposition principle.[8,9] If it is assumed that a step voltage V is applied to a dielectric at $t=0$, and that the absorption current follows the law

$$i_a(t) = VCDt^{-n} \tag{1}$$

Paper 55-456, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 28, 1955; made available for printing April 19, 1955.

C. W. Ross and E. B. Curdts are with the James G. Biddle Company, Philadelphia, Pa.

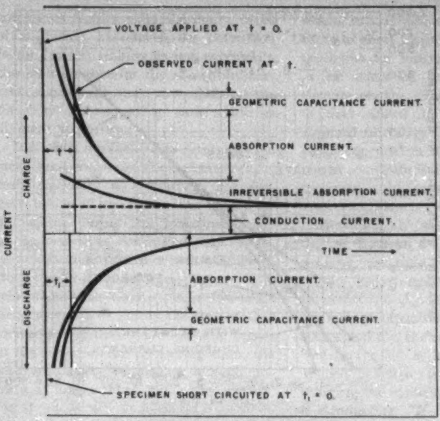

VOLTAGE APPLIED AT t = 0.
OBSERVED CURRENT AT t.
GEOMETRIC CAPACITANCE CURRENT.
t
ABSORPTION CURRENT.
IRREVERSIBLE ABSORPTION CURRENT.
CONDUCTION CURRENT.
TIME
t_1
ABSORPTION CURRENT.
GEOMETRIC CAPACITANCE CURRENT.
SPECIMEN SHORT CIRCUITED AT $t_1 = 0$.
CURRENT
CHARGE
DISCHARGE

Fig. 1. Charge and discharge current components, and their variation with time

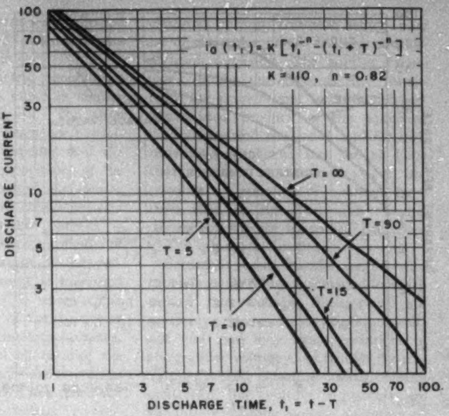


Fig. 2. Theoretical discharge currents for various charging times

then the discharge current $i_a(t_1)$ obtained by applying a short circuit at some time $t=T$ would be given by

$$i_a(t_1)=VK(t_1^{-n}-(t_1+T)^{-n}) \qquad (2)$$

where

$t_1=t-T$ and $K=CD$

The theoretical discharge currents for $K=110$, and $n=0.82$ are plotted in Fig. 2 for various charging times T. Experimental discharge current curves are shown in Fig. 3, for a polyester-impregnated mica tape bar. The effect of insufficient charge time on the discharge current is to give a value of n which is too large, and a value of K which is too small. Also, specific absorption would be less than the actual value. The curves for incomplete charging times deviate quickly from a straight line on a log-log plot as the charge time is shortened.

If correct values are to be obtained for K and n, the charge time must be sufficiently long so that the absorption current at the end of the charge period is negligible compared to the discharge current during the time where significant measurements are made. This condition is approached in the 90-minute curve, Fig. 3. The discharge current in this case plots as a straight line for a considerable time, then shows a gradual break and later follows a straight line of a different slope. It is interesting to note that the slope of the curve for longer times is approximately $2-n$. This could indicate incorrectly that the absorption current follows the law given by Cole and Cole,[10] and has a relaxation time determined approximately by the

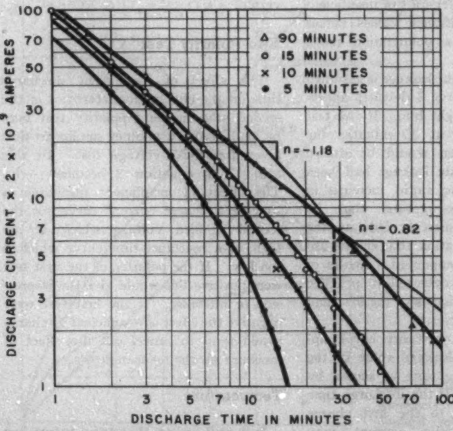


Fig. 3. Experimental discharge currents for various charging times

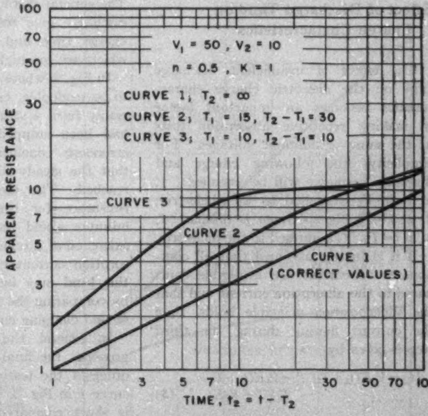


Fig. 4. Effect of various discharge times on charge characteristics: absorption current only

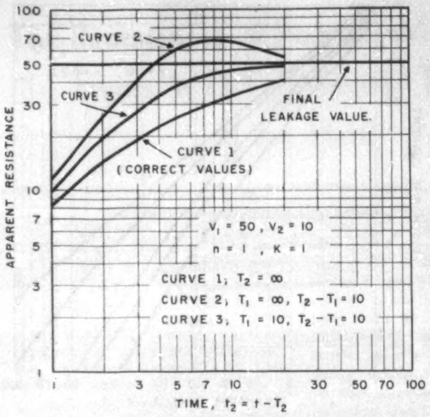


Fig. 5. Effect of various discharge times on charge characteristics, a leakage component included

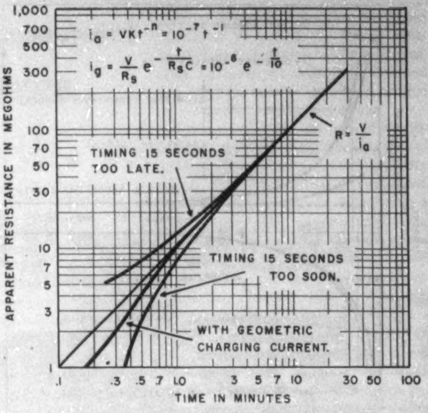


Fig. 6. Dielectric absorption test errors due to incorrect timing and to significant geometric charging current

intercept of the two straight lines. Actually the change in slope in this case does not necessarily indicate a relaxation time as might be supposed, but is due largely to insufficient charging time.

It should also be pointed out that the charge current is not always a reliable indication of negligible absorption current due to leakage current tending to mask it. Even though an observer may not see an appreciable change in current for a rather long period, the specimen may still be charging.

Effect of Discharge Time on Charge Characteristics

The effect of insufficient discharge time on the dielectric charge characteristics becomes an important factor in making repetitive absorption tests at the same or different voltages. For simplicity, the following charge and discharge sequence will be assumed: 1. at $t=0$, V_1 is applied to the dielectric; 2. at $t=T_1$, the specimen is discharged; 3. at $t=T_2$, V_2 is applied to the dielectric.

If it is further assumed that all components of current are negligible compared to the absorption current and that the superposition principle holds, then the current flowing during the third step is given by

$$i_{a3}(t_2) = V_1K(t_2+T_2)^{-n} - V_1K(t_2+T_2-T_1)^{-n} + V_2Kt_2^{-n} \quad (3)$$

where

$t_2 = t - T_2$

A variety of curve shapes may be found for the second charge current, depending on the value of n for the dielectric, the ratio of the applied voltages V_2 to V_1, the first charge period T, and the discharge time T_2-T_1. The number of curve shapes possible is further increased when the leakage current is added to equation 3. From a review of a large number of 2-voltage dielectric absorption tests it was found that curve shapes resulting from insufficient discharge time could be exemplified by the theoretical curves given in Figs. 4 and 5. The general characteristics of insufficient discharge give resistance values versus charge time and a polarization index which are too high.

In Fig. 4, where the leakage is assumed to be negligible, curve 3 deviates appreciably from a straight line. If the test had been stopped at 30 minutes the erroneous conclusion would be drawn that the steady-state leakage had been reached. The more rapid increase in resistance for times greater than 30 minutes would incorrectly indicate the existence of a relaxation time in the absorption current component. Errors of this kind may be detected in practice by comparing the shapes of the first and second charging curves.

In general, the resistance curve approaches the final leakage value in the order of the testing time as shown by curve 1 in Fig. 5. If the discharge time is short compared to the first charge period the absorption current may actually become negative and the resistance-versus-time curve will actually overshoot the final leakage value, as shown by curve 2. Many errors due to insufficient discharge time are not so obvious from the shape of the curve as shown by curve 3 in Fig. 5. The necessary discharge time required to reduce the magnitude of the discharge current to a negligible value is shortened as the ratio of the first voltage step to the second is reduced. For example, the errors shown in Figs. 4 and 5 would be reduced by 25 times if the magnitudes of the two voltages had been reversed.

Two-Polarity Test

The effect of insufficient discharge time on the charge characteristics of the second step in the 2-polarity test may be analyzed in a manner similar to that given for the 2-voltage test. In this case V_2 in equation 3 becomes $-V_1$. The effect of insufficient discharge is to give a charge current curve for the second applied voltage which is too high, or a resistance-time curve which is too low. If the polarity of the first test were positive this would give the impression of moisture. If the first test were negative the effect of insufficient discharge would tend to cancel out the effect of moisture on the measurements.

Temperature

It is well known that temperature has a pronounced effect on resistance and correction factors exist for many types of insulation. However, the correction factor for the leakage is not necessarily the same as that for the absorption current component and may add to the difficulty of comparing data made at different temperatures.

It is often impractical to determine the mean temperature of the dielectric under test, especially if a machine has just been taken off the line. The cooling of a machine during test sometimes leads to the erroneous conclusion that insulation resistance increases with voltage in the two-or-more-voltage test. The relaxation time when present for a dielectric occurs earlier as the temperature increases. If this relaxation time occurs in the order of the measuring time, the resistance curve may show a bend upward at longer times, similar to curve 3, Fig. 4.

Timing and Geometric Charging Current

Incorrect timing for the start of a dielectric test and the geometric charging current component become serious sources of error when significant measurements are attempted at short times. If the applied voltage or short circuit is applied instantaneously and the geometric charging current is negligible, the error due to incorrect timing is generally easy to detect, as shown by the upper and lower curves in Fig. 6. The timing error results in a deviation from a straight-line curve at short times and tends to become negligible for longer times.

The effect of geometric capacitance current on the charge curve is similar to that shown in Fig. 6. The time required to charge the geometric capacitance depends on the impedance of the power source and the capacitance of the specimen. In most cases the power source impedance may be considered as a resistor R_s. If the direct current is derived from an a-c source by means of a diode, the time constant of the geometric capacitance current is increased to about 3.6 R_s $(C+C_s)$, where C_s is the output capacitance of the d-c power source and C is the specimen capacitance. The factor 3.6 is involved because the applied voltage is only effective in charging the capacitance when the diode is conducting. If the voltage is controlled by means of a variable voltage autotransformer, or induction regulator, it is difficult to determine the effective starting time. This also means that a true step voltage is not applied, which further affects the charging time. Rather than make the complex corrections for these factors, only those measurements taken after these factors are negligible are generally considered as significant.

Line Voltage Fluctuations

Line voltage transients, even of the smallest magnitude, cause a geometric charging current transient which may be of the order of the leakage and absorption currents. The pointer excursions of the current-measuring instrument are proportional to the resistance-capacitance product of the specimen capacitance and resistance, and for high resistivity specimens may be so large that accurate measurements are impossible. A number of methods have been used in the field with varying degrees of success to reduce the effect of line voltage fluctuations.

Saturable core transformers having rapid response times are available which reduce line voltage fluctuations appreciably. However, varying input voltage and load may cause variations in the rectifier voltage even though the rms voltage is held constant. Electronically controlled line voltage regulators are capable of holding voltage variation to 0.01 per cent, but their response time is somewhat longer than the saturable core type of regulating transformers. A combination of these methods improves the line voltage stability to a satisfactory degree in some cases. Motor-generator combinations using synchronous motors may improve the transient voltage stability because of the voltage-speed stability of such motors and because of the flywheel effect of the rotating mass. However, it has been found that the use of this principle, although sometimes helpful, is not adequate for higher resistance-capacitance products.

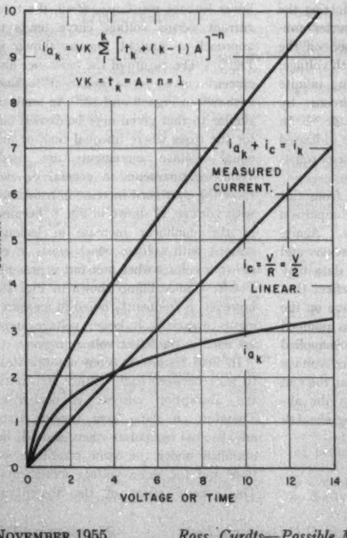


Fig. 7 (left). Currents versus applied voltage for linear leakage

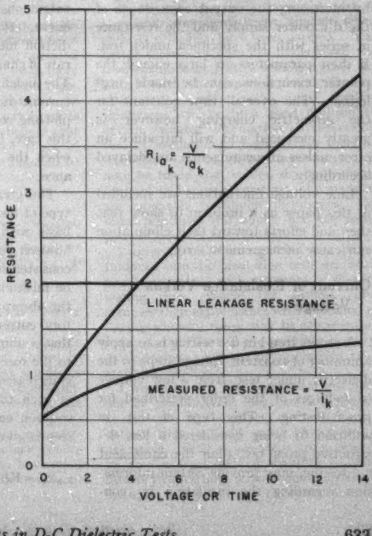


Fig. 8 (right). Resistance versus applied voltage for currents in Fig. 7

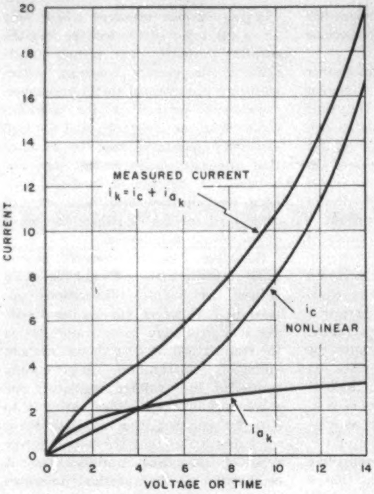


Fig. 9 (left). Currents versus applied voltage for nonlinear leakage

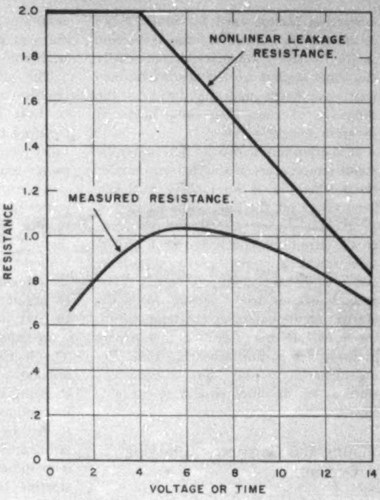


Fig. 10 (right). Resistance versus applied voltage for currents in Fig. 9. Absorption component, same as that in Fig. 8, not shown

Depending somewhat upon the ballistic characteristics of the current-measuring instrument used, the magnitude of the pointer excursions may be reduced by increasing the damping of the instrument by means of a shunt capacitance. However, the slower response of the instrument may make it impossible to recognize the true average reading. The current fluctuations may also be reduced by increasing the output capacitance of the d-c power supply and the resistance in series with the specimen under test. If these parameters are large enough the pointer excursions can be made negligible. The over-all time constant for the geometric charging, however, is greatly increased and will introduce an error unless measurements are delayed accordingly.

Line voltage fluctuations are included in this paper as a problem to show that they and efforts toward their elimination can cause measurement errors.

Current or Resistance Versus Voltage

A recent trend in d-c testing is to apply a number of constant voltage steps to the dielectric under test, from a low voltage to voltages of the order prescribed for proof testing. This type of test, in addition to being considered a less destructive proof test than the equivalent a-c testing, may provide certain information regarding the insulation by considering the current or resistance versus the applied voltage. When the current measured at the end of each voltage step is plotted versus the voltage there may be a disproportional increase in current. In some cases the actual breakdown voltage may be indicated by extrapolating the current curve to infinity.[5] The prediction of the probable breakdown voltage depends to a large extent on the scale used in plotting the curve. It appears that a better prediction method would be based on the rate of change of the current with voltage. The indefiniteness of predicting infinite current is to an extent overcome by plotting resistance versus voltage.[4,6] In this case, breakdown would be indicated when the curve approaches zero resistance.

The greatest value of data from this type of test is probably on a comparison basis with previous test data. Again, however, unless certain precautions and consistencies are followed the data may be misleading because of the effect that the absorption current may have on the total current measured. If it is assumed that a staircase type of voltage is applied to the specimen, e.g., the applied voltage is increased by the amount V at the end of each time interval A, then the absorption current flowing during the kth step is given by

$$i_{ak}(t_k) = VK \sum_{1}^{k} (t_k + (k-1)A)^{-n} \qquad (4)$$

A plot of the current which would be measured at the end of each voltage step for unity values of the parameters in equation 4 is given in Fig. 7, and the equivalent resistance curve is shown in Fig. 8. The absorption current curve tends to approach a steady-state value and the equivalent apparent resistance tends to increase with each step voltage. If the dielectric is assumed to have a linear leakage resistance, $R=2$, the total current versus voltage curve tends to approach a straight line as shown in Fig. 7. The nature of the resistance and current curves for a nonlinear leakage is shown in Figs. 9 and 10. An analysis similar to that given may be carried out for the cases where unequal time or unequal voltage increments are used. Field measurements in general do not show the continued increase in resistance with voltage, as shown in Fig. 8, because of the nonlinear increase in leakage current with voltage which exists, even at lower voltage when moisture is present. The resistance hump shown in Fig. 10, however, is frequently noted in measurements especially if larger voltage steps are used in the lower voltage region.

In field tests it is often not practical to wait between each voltage step until the absorption current is negligible. Therefore, if data from various tests are to be compared they should be obtained under the same conditions of time and the same voltage increments. The general effect of the absorption current is to make the measured resistance appear too low. This effect is most pronounced at the lower voltages. As the number of voltage steps is increased, the measured resistance curve approaches the true leakage value. In cases where breakdown predictions based on leakage are possible the predicted value would be too high if an attempt were made to extrapolate the measured resistance curve in Fig. 10.

Conduction Current at High Voltage

The conduction current components generally behave as shown in Fig. 1 when measured at low voltage. At higher voltages the current may show variations with time even after the absorption current becomes negligible. Such variations are not to be confused with those due to changes in the supply voltage. If the current shows a steady and rather rapid increase with time, the indications are that degradation is taking place and that an incipient fault is developing. Sometimes erratic current fluctuations are noted, which indicate discharges somewhere in the test circuit or specimen. These discharges may or may not indicate damage to the insulation.

The source of one of the more serious pitfalls encountered at high voltage is corona which forms outside of the dielectric under test. Unrestricted corona in air often takes place at terminals, bushings, disconnect switches, and the connecting leads. At constant voltage the corona current may appear as a steady reading on the ammeter. The magnitude increases nonlinearly with the applied voltage and may be appreciable compared to the leakage current of the insulation, even on large equipment. These corona currents would indicate erroneously a poor condition of the insulation and an incorrect breakdown voltage for the dielectric.

Conclusions

Important information may be obtained from d-c tests which is related to the condition of the insulation. That valid interpretations may be made from d-c tests, regarding the condition of insulation, it is necessary that the theory of the tests be well understood and that certain precautions be followed in making measurements.

It is not difficult to show that many field practices lead to erroneous results. These errors are not obvious when the measurements are being made, but they can cause serious mistakes in interpretation.

None of the precautions required in making accurate d-c measurements in the field, as discussed in this paper, should be considered as impractical. Accurate measurements can and are being made, and efforts to abide by these precautions are worth while, since no tests can be justified unless valid interpretations are made from the data obtained.

References

1. The Basis for Nondestructive Testing of Insulation, **R. F. Field.** *AIEE Transactions (Electrical Engineering)*, vol. 60, Sept. 1941, pp. 890–95.

2. Interpretations of Current-Time Curves as Applied to Insulation Testing, **R. F. Field.** General Radio Company, New York, N. Y., 1944.

3. Dielectric Fatigue of Insulating Materials and the Maintenance of Electrical Apparatus, **R. Langlois-Berthelot.** *Bulletin No. 105*, CIGRE, Paris, France, 1954, pp. 14–15.

4. Diagnoses of A-C Generator Insulation Condition by Nondestructive Tests, **A. W. W. Cameron.** *AIEE Transactions*, vol. 71, pt. III, Jan. 1952, pp. 263–74.

5. Testing Electrical Insulation of Rotating Machinery with High-Voltage Direct Current. **G. Leslie Hill.** *Ibid.*, vol. 72, pt. III, Apr. 1953, pp. 159–74.

6. Techniques and Examples of High-Voltage D-C Testing of Rotating-Machine Windings, **C. L. Sidway, B. R. Loxley.** *Ibid.*, pp. 1121–29.

7. The Characteristics of Insulation Resistance, **S. Evershed.** Institution of Electrical Engineers, London, England, vol. 52, no. 224, 1913.

8. **J. Curie.** *Annales de Chimie et de Physique.* Paris, France, vol. 6, no. 18, 1889, p. 203.

9. **J. Hopkinson.** *Philosophical Magazine*, London, England, vol. 5, no. 18, 1876, p. 314.

10. Dispersions and Absorptions in Dielectrics, Pt. II, Direct Current Characteristics, **K. S. Cole, R. H. Cole.** *Journal of Chemical Physics*, vol. 10, Feb. 1942, pp. 98–105.

◆

Discussion

F. S. Oliver (Doble Engineering Company, Belmont, Mass.): The authors should be highly commended for frankly discussing sources of difficulty that may arise when making d-c tests in the field. The time required is one reason for a-c dielectric loss and power factor testing at nondestructive voltages being the most widely accepted test method for most types of high-voltage apparatus insulation as shown by the AIEE questionnaire reported in 1951.[1]

We should like some amplification of the statement that there is a trend towards d-c field testing of insulation. Admittedly, high-potential d-c testing of cables has been well established, with more and more high-potential testing of rotating machinery in recent years. We do not see any such definite trend on other insulations; in fact our experience indicates that there is an increasing trend towards the use of a-c dielectric loss and power factor tests for maintenance testing of most types of insulation in the field.

Also, the statement that the d-c test set is much more portable than an equivalent a-c test set requires amplification. If "equivalent" means equivalence as to voltage in the case of high-voltage test sets, then the statement is true. On the other hand, if "equivalent" indicates the relative value of test results, there are instances in the literature which show that a relatively small a-c test equipment operated with the proper technique can yield more significant results than a much larger high-voltage d-c test set. It should not be construed from these remarks that d-c tests are of little value. Rather, we believe that there are applications for both types of tests, but the capabilities of each should be recognized.

Reference

1. Results of a Questionnaire Covering Current Practices in Electrical Tests on Dielectrics in the Field, AIEE Committee Report. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 986–89.

A. W. W. Cameron (Hydro-Electric Power Commission of Ontario, Toronto, Ont., Canada): The paper is particularly valuable and timely, in view of current interest in high d-c testing. The apparent simplicity of d-c measurements may easily be misleading. Measurements obtained without the precautions so clearly explained in the paper will probably be inconsistent, accounting perhaps for the unsatisfactory results reported by some of those who have tried to avail themselves of the benefits offered by d-c insulation testing.

On the other hand, d-c insulation testing is neither complicated nor difficult when once its principles have been grasped. The better these principles are understood, the more significance will be obtained from the tests. With the paper at hand, the experimenter armed with quite modest equipment will be in a good position to obtain valuable information on the condition of insulation in high-voltage rotating machines. As experience accumulates, it may be found that cables, dry-type transformers, and some other types of apparatus are amenable to similar techniques; the same principles will apply, and the same precautions will be required.

I agree that the errors which may be introduced into d-c insulation tests are commonly associated with attempts to shorten testing time. Would the authors comment on their experiments with a method of testing whereby times may be considerably reduced, the effect of the resulting cumulative adsorption current being clarified by a graphical method?

I would remark that a practical test need not be prohibitively long, as is suggested at the start of the paper, even when each voltage step is held until absorption is negligible. Only a few steps are necessary to give a reasonable probability of nondestructive warning of inadequate dielectric strength. The test can be commenced at half or more of the maximum agreed test voltage, and voltage increments made smaller as that maximum is approached. At the smaller voltage increments, times may be somewhat reduced without significant error. The ground insulation of three or more phases or other portions of a winding may be tested simultaneously, and measured individually by suitable switching. I stress this type of test because of a suspicion that longer energization at the higher voltages may be conducive to the establishment of the nondestructive but observable currents in cracks or deteriorated areas, on which nondestructive warning of low strength depends. The behavior of such leakage currents is as yet little understood.

In my organization high-voltage d-c tests of this latter type are scheduled during major overhauls of water-wheel generators and other rotating machines, at times when testing of the stator winding will not interfere with work that is proceeding on the turbine, the removed generator rotor, or other parts at a safe distance. Four hours normally suffice for a test. The organization and effort involved are very small in comparison with the value of the information obtained.

C. A. Duke (Tennessee Valley Authority, Chattanooga, Tenn.): As the advantages of d-c testing are realized through greater use, this paper will contribute greatly to the avoidance of mistakes in the accumulation of data, especially in the setting up of initial procedures.

The effect of previous discharge time on charge characteristics is difficult to evaluate, and this effect will frequently be minimized until its importance is realized. It is much better to allow for this effect before data are collected than to try to explain why the apparent resistance of the insulation of a large generator reaches a value greater than infinity, as is possible if the discharge time previous to testing is too short.

The underlying theme in the paper is to evolve a practical system of d-c testing of insulation with a minimum of deviations which might lead to inconclusive results or to the collection of faulty data which in turn might lead to unjustifiable discrediting of a valuable testing tool.

John C. Parker (Memphis Light, Gas and Water Division, Memphis, Tenn.): This interesting paper states that: "Important information may be obtained from the d-c tests which is related to the condition of the insulation. Periodic tests make it possible to follow the insulation deterioration, and thereby aid in detecting the need of reconditioning steps and estimating life expectancy." It is assumed that this statement applied to the so-called nondestructive d-c tests, i.e., d-c tests made with potentials low enough to ensure against failure of insulation regardless of its condition.

Although this test is most useful in detecting and following some types of insulation deterioration and contamination, it will not detect other types of insulation deterioration. For example, a d-c test will not detect corona-cutting in otherwise good rubber insulation unless the d-c test voltage is sufficiently high to puncture the defect.

Recognition of the fact that d-c tests are very useful for detecting and following some types of insulation deterioration and contamination and are absolutely useless for detecting other types, and recognition of the possible measurement errors, which are so well described by the authors, will greatly increase the usefulness of this test.

Graham Lee Moses (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The authors are to be congratulated on their objective, and possibly not very popular, view of an important subject. They are practical theorists and have clearly described and segregated the components of a complex problem which is too generally oversimplified. Most scientific investigators of d-c testing are familiar with the theories and problems described, but many have not been well aware of the facts and theoretical background. As stated, precautions in making measurements are not impracticable, but necessary to obtain accurate data for valid interpretation.

J. S. Johnson (Westinghouse Electric Corporation, East Pittsburgh, Pa.): This is an interesting and informative paper. High-voltage d-c testing, particularly in the evaluation of the condition of large rotating machine insulation, is enjoying increasing acceptance. It is quite obvious that the use of direct current in the evaluation of insulation condition is very helpful and provides more information than is obtainable by other means. For results to be most meaningful, careful attention should be direct toward the use of proper test methods and the recognition of variables resulting from the manner in which a test is applied, rather than from some property of the insulation itself.

Figs. 4 and 5 clearly demonstrate the effect of various discharge times on the dielectric absorption characteristics observed at several voltages. Since it is generally desired to note the effect of test voltage on the absorption characteristics, the effects of discharge time should, obviously, be minimized. On the other hand, time is usually of the essence in the conduct of these tests and it becomes undesirable to prohibitively long discharge times. K. M. Stevens and I have reported the results of a series of dielectric absorption characteristics made at 5-kv increments starting from 5,000 volts direct current.[1] In these tests the duration of test voltage for each step was 10 minutes and the discharge time $^1/_2$ hour. The results are illustrated in Fig. 1 and Table I. It is apparent from the examination of these data that the $^1/_2$-hour discharge time was sufficient to discharge the winding for the machine in question. It is, of course, important always to make the absorption tests in increasing order to test voltage if prohibitively long discharge times are to be avoided. In dielectric absorption tests made on a relatively large number of high-voltage stator windings in which 10-minute absorption tests were run at 5, 10, and 20 kv respectively, it was also found that $^1/_4$ hour in most cases is satisfactory to minimize the effects of residual charge.

The authors' cautions concerning the effect of corona on lead wires to the equipment under test are well taken.

Recognition of the measurement pitfalls that can be encountered in d-c dielectric testing is helpful in the obtaining of meaningful data. In some cases, depending upon interpretation of the data and the time element involved in obtaining the information, some short cuts are necessary and perhaps even desirable. However, when this is done, it should always be done with a clear understanding of the effect on the measurement characteristics.

REFERENCE

1. DESTRUCTIVE A-C AND D-C TESTS ON TWO LARGE TURBINE GENERATORS OF THE SOUTHERN CALIFORNIA EDISON COMPANY, **K. M. Stevens, J. S. Johnson.** *AIEE Transactions*, vol. 73, pt. III-B, Oct. 1954, pp. 1115-26.

C. W. Ross and **E. B. Curdts:** The authors are gratified that the practical value of the paper is appreciated by those who contributed discussions. In amplification of the statement mentioned by Mr. Oliver concerning the trend toward d-c field testing of insulation, from an application engineering and sales point of view, our company is experiencing an increasing interest in and increased sales of d-c test sets for preventative maintenance testing for such equipment as cable, rotating machinery, and switchgear, particularly for tests at voltages in excess of the service voltage rating of the equipment to be tested.

As to the relative portability of d-c and a-c test equipment, the word "equivalent" was primarily intended to mean equivalence as to the test voltage applied, which Mr. Oliver concedes to be true. However, since the issue has been raised, we can also show evidence where in some cases smaller d-c test equipment operated with the proper technique can yield significant results, at least equivalent to larger high-voltage a-c test equipment. We intended no criticism of economical and technically significant a-c test methods.

Mr. Cameron has suggested that we comment on the method we advocate, for economic reasons in certain cases, which shortens the testing time without necessarily affecting the significance of the kind of test discussed in the section "Current or Resistance Versus Voltage." The successful use of this short-time method presupposes a thorough understanding of equation 4 and the typical behavior patterns shown in Figs. 7 through 10.

Incidentally, it may be noted that equation 4 can be modified from the point of view of mathematical rigor. In its present form the parameter t_k is defined as the variable time within k^{th} steps. This form produces a correct numerical solution, but reverses the order of the terms in the summation. A more rigorous form would be

$$i^a{}_k(t) = VK\sum_1^k [t+(1-k)A]^{-n}$$

where

t = the total time from start of test.

The basic idea of this more economical technique is to use a short fixed time for each voltage step even though the absorption current may be significant during the time interval chosen. This is in contrast to the method commonly used where the intention is to hold voltage steps constant until the absorption current has become relatively insignificant. There are two reasons for this latter technique. One is to avoid the necessity of interpreting the degree of nonlinearity of the conduction current from data which include the time-dependent absorption current component. In other words, it is felt by some investigators that the absorption current distorts the current- or resistance-voltage behavior to the extent that interpretation becomes difficult. We believe that once this behavior is understood it can be used to advantage rather than confusing the issue.

The other reason the complete charge technique is preferred by some is that voltage gradients throughout the insulation are considered not stabilized for test purposes until the absorption phenomenon is essentially complete. Our experiences so far have not given evidence that this possibility constitutes a serious objection to the short fixed-time test method.

We have used the short fixed-time method in the field often enough to satisfy ourselves that considerable time is saved in the test and that significant results are obtained. We are using 1-minute steps, shorter intervals than that being impractical in most cases because of distortion caused by the geometric-capacitance current. We agree with Mr. Cameron, however, that longer voltage steps at the higher voltage levels may possibly bring to light further evidence of nonlinearity or instability. We cope with this possibility, in cases where the test reveals questionable conditions in the preceding voltage steps, by prolonging the final step.

We have also in several field tests broadened the method by using voltage steps of different magnitudes without introducing confusing distortions in the resistance-voltage curve. In this case equation 4 is rewritten so that the constant step-voltage V is made a variable and included in the summation Then

$$i_{a_k}(t) = K\sum_{1}^{k} \Delta V_k [t+(1-k)A]^{-n}$$

where

ΔV_k = step (incremental) voltage applied at the k^{th} step

t = total time from the start of test

A = constant time interval between step voltage applications

The total current becomes $i_t = i_{a_k} + i_c$

For the interpretation of i_c under nonlinear conditions, see reference 5 of the paper.

Fig. 11 shows a typical result of testing an important 2,300-volt motor in a large industrial plant, with three 1-minute steps of $\Delta V_1 = 1{,}000$ volts, $\Delta V = 1{,}500$ volts, and $\Delta V_3 = 2{,}500$ volts. Curve 1 indicates conditions as found in terms of insulation resistance at increasing voltage. It will be noted that nonlinear conduction current masks all evidence of the absorption current, even in the intermediate 2,500-volt ($\Delta V_2 = 1{,}500$ volts) step. If the motor had been in somewhat better, but still questionable, condition the intermediate resistance value may have been higher than the first step value, similar to that shown in Fig. 10, because of a greater relative magnitude of absorption current at that voltage. Curve 2 indicates conditions as found for the same winding, after a cleaning, baking, and impregnating operation. Here, both the intermediate and final voltage steps show a rising tendency, similar to that given in Fig. 8, indicating the practical elimination of conduction current nonlinearities, borne out by revealing the presence of a significant absorption current component.

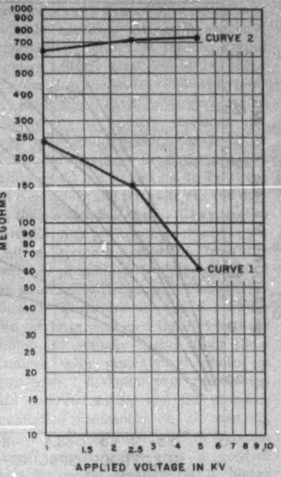


Fig. 11. **Resistance versus applied voltage curves for a large motor before and after reconditioning**

The Dielectric Behavior of Some Fluorogases and Their Mixtures

G. CAMILLI
FELLOW AIEE

T. W. LIAO
MEMBER AIEE

R. E. PLUMP
NONMEMBER AIEE

IN THE last 15 years the practical possibilities of the fluorogases as dielectrics have increased and their use has been demonstrated in several applications.[1,2] Sulfur hexafluoride and Freon 12 are available economically in large quantities and for this reason have received extensive studies.[3,4,5] However, fragmentary data showed that many compounds, although not yet available in large quantities, held promise and needed comprehensive evaluations.

The corona starting voltage and the impulse and 60-cycle strength of the new fluorogases have been investigated in a uniform and in a nonuniform field and comparative values with both SF_6 and nitrogen are given in this paper. Because of the well-known "anomalous behavior" of electronegative gases, the strength of these fluorogases in a nonuniform field begin to drop in a certain region of pressures. It was found possible to shift this region for practical purposes to a higher pressure range by mixing the gases with nitrogen. Additional benefits of this procedure were realized also in lowering total gas cost and condensation temperature which is very desirable for outdoor applications. These advantages resulting from mixing nitrogen should find radical applications.

Test Setup and Procedure

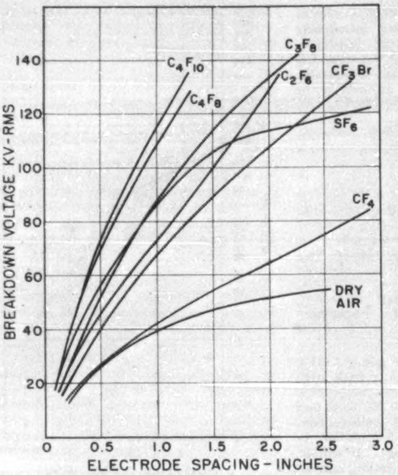
The test cell made of a Plexiglas cylinder of 8-inch inside diameter with herkolite end plates was used. Two sets of test electrodes were chosen to simulate the fairly uniform field and the nonuniform field conditions often encountered in an electrical apparatus. The former consists of 3-inch-diameter chrome-plated planes with rounded edges. The other set consists of a 1/4-inch square aluminum rod and a 3-inch-diameter grounded plane. The electrode spacing was ad-

---

Paper 55-440, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 14, 1955; made available for printing April 8, 1955.

G. CAMILLI, T. W. LIAO, and R. E. PLUMP are with the General Electric Company, Pittsfield, Mass.

The authors wish to express their indebtedness to J. H. Hagenguth and L. Wetherill for their interest and encouragement throughout the progress of this work, and to A. Harju and C. B. Musick for their assistance in obtaining the experimental results.



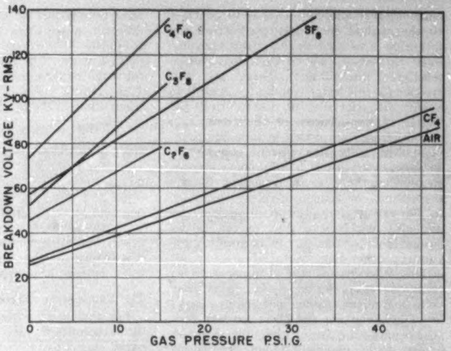


Fig. 1 (left). 60-cycle strength of fluorogases, 3-inch-diameter round-edged plane to plane, atmospheric pressure

Fig. 3 (above). Dielectric strength of fluorogases versus pressure, 3-inch-diameter round-edged plane to plane, 1/2-inch spacing

justed by means of a micrometer connected to one of the electrodes.

Before each fluorogas was admitted into the cell, the cylinder was first washed with soap and water, rinsed with distilled water, and dried in an oven. After the cylinder was assembled in place, the cell was evacuated two times between which it was flushed with the test gas. Then the test gas was admitted to the desired pressure for test.

During the 60-cycle breakdown voltage tests the voltage was raised at a rate of about 1 kv per second until breakdown occurred. For corona tests the voltage was raised at a slower rate but in neither case was the rate of rise found to be critical. During impulse tests the voltage was increased in approximately 10-per-cent steps and one 1.5x50-microsecond impulse was applied at each step until breakdown occurred. Three breakdown values were taken for each gap spacing for either type of voltages.

Dielectric Strength of Fluorogases

The dielectric strength of fluorogases is, in general, much higher than that of air. The breakdown voltages versus electrode spacing with a pair of 3-inch-diameter round-edged planes are plotted in Fig. 1.

It is seen that the slopes of the curves are not alike for the various gases, especially at large spacings. This is attributed to the deviation from uniform field at the electrode edges for such spacings, and the different reactions of the gases to the field distortion. Therefore, it is not proper to compare the relative intrinsic strength of the various gases at those spacings. However, at small spacings where the field is fairly uniform, a comparison of the relative strengths can be made. By plotting the breakdown voltages of the various gases at 1/2-inch spacing versus their molecular weights, it can be concluded that the dielectric strength of the gases, as shown in Fig. 2(A), increases roughly with the molecular weight. Similar relationship was also found in hydrocarbon liquids where the dielectric strength increases with density[6] and with molecular chain length.[7] It is interesting to note that the nonuniform field strength of the gases seem to increase even faster with the molecular weight, as shown by curve *B* of the same figure. However, this relationship should not be generalized at this time.

The dielectric strength of these gases have also been determined at pressures higher than atmospheric. The linear increase with pressure, shown in Fig. 3, is in accord with Paschen's law. However, the small difference in slope for the various gases is not clear. This pressure dependence has been utilized in gas-insulated apparatus[1,2,8] in which the elec-

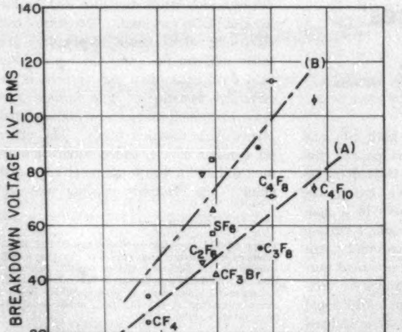


Fig. 2. Correlation between dielectric strength and molecular weight of fluorogases

A—3-inch-diameter round-edged planes spaced 0.5 inch apart

B—1/4-inch square rod to 3-inch-diameter plane spaced 1 inch apart

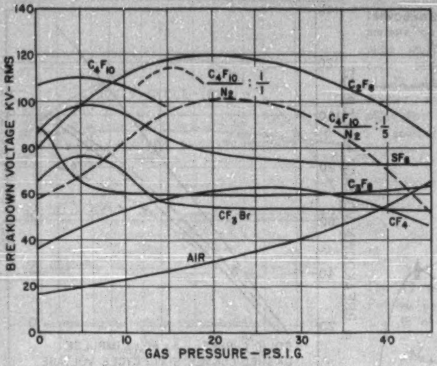


Fig. 4. Dielectric strength of fluorogases versus pressure, 1/4-inch square rod to plane, 1-inch spacing

trode configurations give essentially uniform fields. However, it is well known that in nonuniform fields the strength of the electronegative gases actually decreases with increasing pressure in some region of pressures.[9] This anomalous behavior is possessed by all electronegative gases or a mixture of gases that contain an electronegative gas such as O_2 in air.[10] For a 1-inch rod-to-plane gap, the anomalous behavior of some electronegative fluorogases is shown in Fig. 4. It is seen that the critical pressure at which the strength begins to drop is different for different gases.

For a specific gas this anomaly can also be observed by plotting the strength of the gas versus electrode spacing, as shown in Fig. 5. It is seen that this anomaly occurs when the breakdown curves deviate from a linear increase. For instance, for pressures of 15 and 45 pounds per square inch gauge, the anomaly starts at about 0.5- to 1-inch spacing. At atmospheric pressure, the anomaly does not occur within a spacing range of a few inches for most of the gases shown in Fig. 6.

When the curves in Fig. 6 are compared with those of Fig. 1, it is found that the strengths of these gases determined with the rod-to-plane and with the parallel planes are comparable except for dry air. This is caused by the fact that the breakdown of the rod-to-plane gap is preceded by corona. By virtue of the electronegative property of these fluorogases, corona produces a much larger cloud of negative ions than in the case of dry air. When the rod is negative, these ions minimize the high gradient at the rod resulting in a better uniformity of the field. Consequently, the breakdown values are close to the ideal parallel plane values. When the rod is positive, the breakdown phenomenon can be treated in two regions of electrode spacing. At small spacings, negative ions formed at the rod during negative voltage crest would be swept into the plane before the following positive half-cycle. Thus, breakdown can be initiated by a positive streamer. At large spacings, however, negative ion formed during almost the entire negative half-cycle do not have time to be swept out completely and return to the rod during the positive half-cycle. They tend to neutralize any positive ions created by a positive streamer emanating from the rod, thus prohibiting the propagation of the streamer to the plane. This is schematically illustrated in Fig. 7. Therefore, at large spacings the rod-to-plane values are actually higher than those obtained with round-edged planes in most cases. A similar phenomenon was also observed in air and the streamer prohibition mechanism in air was described in considerable detail in a corona paper[11] which states that a positive streamer would emanate at a higher voltage from a smaller wire than from a large diameter one. This phenomenon often appears contrary to the common experience in that rounded electrodes should result in a higher breakdown voltage than sharp electrodes.

The visual corona starting voltages for the various gases are shown by the dotted curves of Fig. 6. To avoid inconsistency of human eyes during the tests, a photo cell incorporated with an amplifier and an indicating meter was used. A light detection method was preferred because there was some minute corona in the test cell which confused the indication by conventional electrical detection methods.[12,13] It is conceivable that these corona

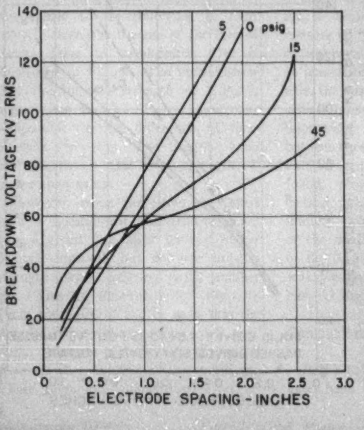
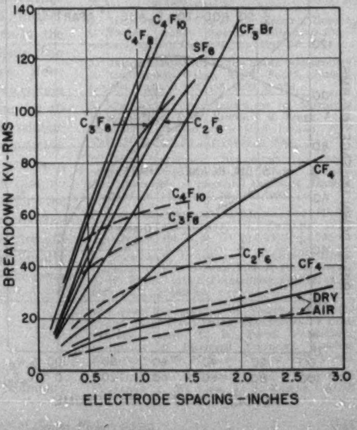


Fig. 5 (left). 60-cycle strength of CF_3Br with 1/4-inch square rod-to-plane gap at several pressures

Fig. 6 (right). 60-cycle corona and breakdown voltage of fluorogases, 1/4-inch square rod to plane, atmospheric pressure

Solid curves—Breakdown
Dashed curves—Corona



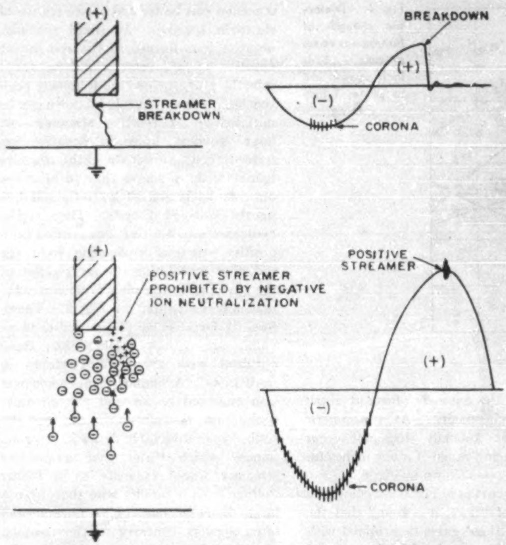


Fig. 7 (left). Schematic diagram showing ion distribution in rod-to-plane gaps

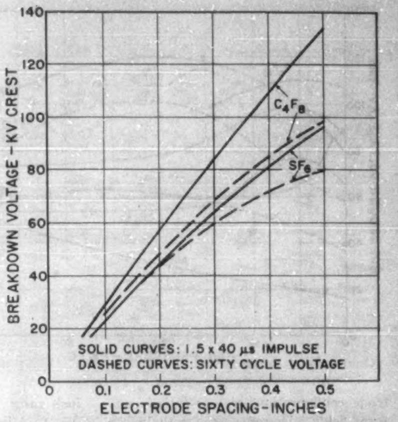


Fig. 9 (above). Comparison of impulse and 60-cycle breakdown characteristics of fluorogases, 3-inch-diameter round-edged plane to plane, atmospheric pressure

values might be slightly higher than the actual corona starting voltages. The wide separation of corona and breakdown values are often referred to as due to "corona stabilization."

Dielectric Strength of Mixtures of Fluorogases and Nitrogen

In searching a means to avoid the anomalous behavior of the fluorogases with nonuniform field, nitrogen was mixed with C_4F_{10} in various proportions. It was extremely interesting to find that at 15 pounds per square inch gauge, while the breakdown voltage of pure C_4F_{10} had dropped below 100 kv, Fig. 4, a 1-to-1 mixture actually yielded a higher strength. This is also shown by the solid curve in Fig. 8. It is seen that the strength is maintained above that of the pure C_4F_{10} up to a mixture of 5 (N_2) to 1 (C_4F_{10}). The reason for this is that as N_2 is mixed into C_4F_{10}, the critical pressure of the anomalous region becomes higher. As illustrated by the dotted curves in Fig. 4, a 5 (N_2)-to-1 (C_4F_{10}) mixture will shift the critical pressure to about 25 pounds per square inch gauge

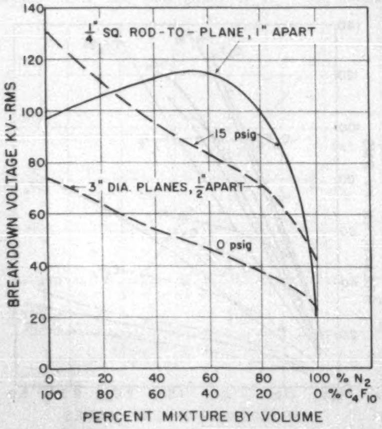
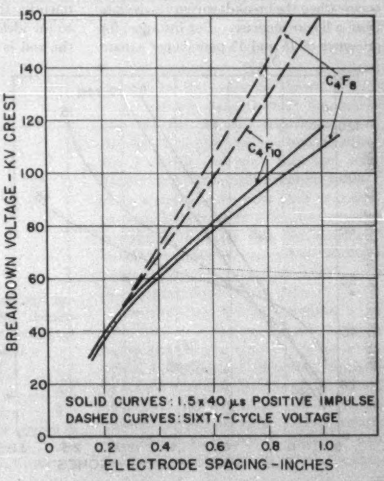


Fig. 8 (left). 60-cycle strength of mixtures of C_4F_{10} and N_2

Fig. 10 (right). Comparison of impulse and 60-cycle breakdown characteristics of fluorogases, 1/4-inch square rod-to-plane gap, atmospheric pressure



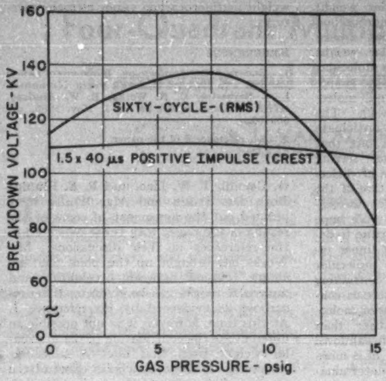


Fig. 11 (left). The effect of pressure on impulse and 60-cycle strengths of C_4F_8, 1/4-inch square rod-to-plane gap, 1-inch spacing

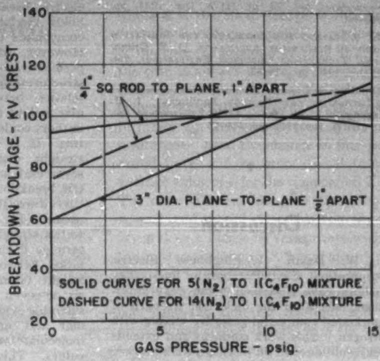


Fig. 12 (right). Positive impulse strength of mixtures of C_4F_{10} and N_2

below which the mixture strength increases with pressure as a normal gas.

In the uniform field the strength of the mixture is lower than that of the pure gas. However, by using higher pressures the uniform field strength for the mixtures is proportionally increased. For instance, at 15 pounds, the breakdown strength of the mixture of 4 (N_2) to 1 (C_4F_{10}) or less is above the pure gas at atmospheric pressure. These are illustrated by the broken curves in Fig. 8.

Impulse Strength of Fluorogas-Nitrogen Mixtures

Data on the impulse breakdown characteristics of the insulating gases are pertinent when they are used in apparatus that is liable to be exposed to lightning or switching surges. It is expected that the difference in impulse strengths of various gases in a quasi-uniform field should be comparable to the difference of their 60-cycle strengths because of the lack of corona prior to breakdown. This is illustrated in Fig. 9. The curves show that the impulse strength of C_4F_{10} is higher than that of SF_6 by an amount comparable to the difference of their 60-cycle strengths. The impulse ratio (impulse to 60-cycle crest strength) is higher than unity.

However, when the gases are tested with a nonuniform field (rod-to-plane) gap, a given gas might have a higher 60-cycle strength but a lower impulse strength when compared with another gas. This is illustrated by Fig. 10. From the figure it can be seen that the impulse ratios are consistently less than unity, except perhaps at the small spacings where the field is less nonuniform. These anomalous characteristics are attributed to the presence and intensity of corona prior to breakdown and the resulting space charge which varies for different conditions. The difference of the corona or space charge effect can be illustrated by plotting the strength of the gas with various pressures, as shown in Fig. 11. It is seen that neither the impluse nor the 60-cycle strength increases consistently with pressure as for an ideal gas. The anomaly is more pronounced in the case of 60 cycles, where the effect of corona or space charge is expected to be larger.

It has been shown previously that the addition of N_2 to the fluorogases effectively shifts the critical pressure to a higher value in the case of the nonuniform field. This was found to be true also in the case of impulse voltages. For instance, in Fig. 12 the nonuniform field strength appears to start dropping slightly at about 10 pounds per square inch gauge for a 5-to-1 mixture. With further increase of N_2 content from 5 to 14, the critical pressure is shifted to higher than 15 pounds per square inch gauge. However, the mixture ratio for a given certain critical pressure is not the same for all fluorogases. As yet, no definite relation between these parameters and the molecular structure of the gases has been found. Similar to the 60-cycle case, no critical pressure was found for the uniform field breakdown characteristics of the mixtures. In fact, the strength of the 5-to-1 mixture increases linearly with pressure, as illustrated by the lower curve of Fig. 12.

Conclusions

In the last few years, many new fluorogases have been available. Preliminary evaluations indicate that in many instances the dielectric characteristics of the new fluorocarbons are superior to those of the earlier halogenated gaseous compounds. In some cases, however, it is desirable to mix these gases with nitrogen for the purposes of lowering the total cost of the gases, to reduce the condensation point of the pure fluorogases, and to obtain a more uniform breakdown characteristic in the nonuniform field.

References

1. A Compact Pressure Insulated Electrostatic X-Ray Generator, **J. G. Trump, R. J. Van de Graaff.** *Physical Review*, New York, N. Y., vol. 55, 1939, pp. 1160–65.

2. Million-Volt X-Ray Outfit, **E. E. Charleton, W. F. Westendorp, L. E. Dempster, G. Hotaling.** *Journal of Applied Physics*, New York, N. Y., vol. 10, 1939, pp. 374–85.

3. The Effect of Pressure on the Positive Point to Plane Discharges in Nitrogen, Oxygen, Sulfur, Hexafluoride, Etc., **H. C. Pollock, F. S. Cooper.** *Physical Review*, New York, N. Y., vol. 56, 1939, p. 170.

4. High-Pressure Gas as a Dielectric, **Gordon C. Nonken.** *AIEE Transactions (Electrical Engineering)*, vol. 60, Dec. 1941, pp. 1017–20.

5. Positive Point to Plane Spark Breakdown of Compressed Gases, **T. R. Foord.** *Nature*, London, England, vol. 166, 1950, p. 688.

6. On the Electric Strengths of Aliphatic Hydrocarbons, **R. W. Crowe, J. K. Bragg, A. H. Sharbaugh.** *Journal of Applied Physics*, New York, N. Y., vol. 25, Mar. 1954, p. 392.

7. The Dielectric Breakdown of Some Simple Organic Liquids, **B. Salvage.** *Proceedings*, Institution of Electrical Engineers, London, England, pt. IV, vol. 1.98, Oct. 1951.

8. An Experimental Gas-Insulated 138-Kv Current Transformer, **G. Camilli.** *AIEE Transactions*, vol. 74, pt. III, Apr. 1955, pp. 100–04.

9. Dielectric Breakdown of Sulfur Hexafluoride in Nonuniform Fields, **C. N. Works, T. W. Dakin.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 682–89.

10. Some Experiments on Positive Point-to-Plane Corona and Spark Breakdown of Compressed Gases, **T. R. Foord.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 100, pt. II, no. 78, Dec. 1953, p. 585.

11. Detection of Corona in Oil at Very High Voltages, **T. W. Liao, J. S. Kresge.** *AIEE*

Transactions, vol. 73, pt. III-B. Dec. 1954, pp. 1389–95.

12. A METHOD FOR DETECTING THE IONIZATION POINT ON ELECTRICAL APPARATUS, G. E. Quinn. *AIEE Transactions* (*Electrical Engineering*), vol. 59, Dec. 1940, pp. 680–82.

13. IMPULSE CORONA DETECTION, MEASUREMENT OF INTENSITY, AND DAMAGE PRODUCED, J. H. Hagenguth, T. W. Liao. *AIEE Transactions*, vol. 71, pt. III, Jan. 1952, pp. 461–65.

◆

Discussion

T. W. Dakin (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The authors have described some very interesting effects on the dielectric strength of fluorinated compound gases by mixing with nitrogen. This work confirms and extends results obtained in the discusser's laboratory. An examination of the authors' data in conjunction with the data in reference 9 of the paper and other unpublished data of the discusser's colleagues indicates that the enhanced dielectric strength of some mixtures of nitrogen and a fluorine compound gas over that of the pure compound at an equivalent pressure is associated with a very nonuniform electric field as in a point or rod gap. This indicates that the effect of the nitrogen in the mixture is to modify the development of the space charge and the streamer development from the pointed electrode.

The authors' interpretation of the a-c breakdown mechanism of these gases emphasizes the effect of the residual negative ions remaining during the positive half-cycle from the corona which occurs on the negative half-cycle. They imply that these negative ions formed during the negative half-cycle have a major effect in reducing streamers on the positive half-cycle. The results of Mr. Works and the discusser shown in Fig. 7 of reference 9 of the paper with a rod-to-plane gap in SF_6 show that the positive d-c and a-c crest breakdown strengths are nearly identical. Thus, the residual negative ions formed during the negative half-cycle can have only a minor effect on the a-c positive cycle breakdown, contrary to the authors' implication. With a rod gap at large spacings corona begins far below breakdown and very early in the positive half-cycle as shown in Fig. 9 of reference 9 (Works and Dakin). This positive point corona produces both positive ions and negative ions in a lesser number. The latter reduce the concentration of positive space charge developing in the positive streamers and increase the dielectric strength.

C. N. Works (Westinghouse Electric Corporation, East Pittsburgh, Pa.): It is pointed out in this paper that the dielectric breakdown of fluorogases is a linear function of molecular weight, as shown in Fig. 2 This is an interesting relationship. The authors tested fluorogases containing from one to four carbon atoms. A recent paper[1] shows that this relationship between dielectric breakdown and molecular weight holds for straight chain perfluorocarbon compounds containing five carbon atoms. However, when higher molecular weight perfluorocarbon vapors of various molecular structures were tested, there was no simple relation between breakdown and molecular weight (Fig. 8 of rererence 1). The authors could not be expected to anticipate this. It is interesting to point out that straight chain perfluoroheptane, having seven carbon atoms, has a value of dielectric breakdown above the extention of the curve through the other straight chain compounds (Fig. 8 of reference 1). We hope to test straight chain perfluorohexane to determine if the departure from a linear relation between breakdown and molecular weight occurs for compounds containing six or seven carbon atoms. There is considerable evidence that the details of molecular and atomic structure rather than molecular weight determine breakdown values. This suggests that work of a more fundamental nature, such as the determination of cross sections for ionization, scattering, and capture, should be done on these fluorogases.

It is interesting that the "critical pressure" at which the dielectric strength falls off is different for different gases as shown in Fig. 4. In a previous paper,[2] we observed that this critical pressure varied with electrode configuration for nonuniform field conditions in the case of another electronegative gas. Thus, it is evident that the breakdown strength of a fluorogas under nonuniform field conditions depends on so many variables that breakdown values must be determined experimentally for a give application.

Fig 7 of this paper explains the long streamers that have been observed to extend from the high-voltage electrode without completely crossing the gap. However, Fig. 3 of reference 1 shows that a maximum of lesser magnitude than for the 60-cycle case occurs in the curve of dielectric breakdown versus pressure at a critical pressure for positive impulse breakdown of perfluorocarbon vapor. This suggests that the space charge around the high-voltage electrode does not develop as much since the time is not sufficient for the migration of ions in the case of impulse voltages.

The curves, Fig. 5, showing breakdown voltage versus electrode spacing for several pressures are similar in appearance to curves obtained some time ago for another electronegative gas, SF_6: see Fig. 6 of reference 2. This provides additional evidence that the performance shown by these curves is characteristic of electronegative gases in general.

The curves of Fig. 11 of this paper showing the effect of pressure on impulse and 60-cycle strength of a fluorogas are in agreement with the curves obtained on another electronegative gas, SF_6, as shown in one of our previous AIEE papers: see Fig. 9 of reference 2.

It is worthy of note that the breakdown strength of mixtures of these fluorogases with nitrogen may be greater than that of a pure fluorogas under nonuniform field conditions. This is in agreement with the results we obtained for higher molecular weight perfluorocarbon vapor-nitrogen mixtures.

REFERENCES

1. ELECTRIC BREAKDOWN OF PERFLUOROCARBON VAPORS AND THEIR MIXTURE WITH NITROGEN, L. J. Berberich, C. N. Works, E. W. Lindsay. *AIEE Transactions*, vol. 74, pt. I, Nov. 1955, pp. 660–66.

2. See reference 9 of the paper.

G. Camilli, T. W. Liao, and R. E. Plump: Both Dr. Dakin and Mr. Works have pointed out the agreement of some of the results in the paper with those indicated in the references in their discussions. Mr. Works has brought up the point that no simple relation between breakdown and molecular weight can be given to all fluorocarbons as evidenced by his reference 1. At this time, however, it is not possible to differentiate the contribution of the molecular weight from that due to molecular structure. Perhaps by a joint effort of the gas manufacturers and the various investigators, we may arrive at a generalized law to predict the dielectric strength of any fluorogases.

Both discussers comment on the interpretation of the a-c breakdown mechanism illustrated in Fig. 7 of the paper. The arguments presented by them are logical. However, the fact that the rod-to-plane gap is weaker than the parallel planes at small spacings and stronger at large spacings is not yet explained. We were trying to explain this breakdown-spacing characteristic rather than that of pressure, although there might be some relation between them.

The emphasis on the residual negative ions was based on a study of the corona phenomenon of a rod or point to plane gap in air. When either positive or negative direct voltage was applied to the sharp electrode, corona started at about the same voltage. However, if 60-cycle alternating voltage was used, then corona was observed only in the negative half-cycle at the same corona starting level. Positive corona did not occur until the voltage was raised several times the negative corona starting level.

In air, only oxygen forms negative ions. It is assumed that negative ion formation in a pure electronegative gas, such as SF_6, may be even greater. The negative ions formed in the positive half-cycle in the case of SF_6 may be sufficient to inhibit the streamer formation, thus the residual negative ions from the negative half-cycle become less effective as stated by Dr. Dakin. Since no mobility data on the various ions (corona decomposition of SF_6) are available and the test cell used in this experiment was not corona-free at the corona levels in SF_6, oscillographic studies of the corona phenomenon were not made. As mentioned in the paper, only a phototube was used for determining the visual corona levels.

We believe that a detailed study of the negative ion space charge in both half-cycles should give an insight into the anomalous phenomena observed under nonuniform field conditions.

Four-Quadrant Multiplication with Transistors and Magnetic Cores

R. L. VAN ALLEN
ASSOCIATE MEMBER AIEE

THIS paper describes a relatively simple 4-quadrant analogue multiplying circuit which may be built with the use of diodes, transistors, and magnetic cores. Preliminary tests on the first experimental model show accuracy of approximately ±1 per cent (%) of full scale. Two circuits are discussed, one for sinusoidal supply and the other for square wave supply. Each provides a product of two independent voltages with output in the form of d-c pulses which may be averaged or integrated to yield a continuous analogue output. The pulsing d-c output is reversible in polarity depending upon the signs of the input, each pulse being width-modulated proportional to one of the input voltages and amplitude-modulated proportional to the other. The compactness and simplicity of the circuit make it especially attractive for applications of control and instrumentation in systems where the operation of multiplication is required.

Background and Theory

In the past, some work has been done with the concept of magnetic amplifiers in computer applications.[1,2] More recently, the introduction of transistors into the magnetic-amplifier field (especially the achievements of Bright[3] and Royer[4]) has increased the scope of magnetic circuitry. Some aspects of each concept have been used to devise the circuit for 4-quadrant multiplication.

Given two independent variables in the form of voltages e_1 and e_2, it is required that these voltages be multiplied and a product obtained in the form of an output voltage. One method, the method used by this circuit, is to produce an output voltage in the form of periodic d-c pulses having width proportional to e_1 and height proportional to e_2, then the area of each pulse will be proportional to the product of e_1 and e_2. To accomplish this, three important criteria for the circuit are that:

1. The pulse width be proportional to e_1.
2. The pulse height be proportional to e_2.
3. The pulse be repetitive.

A simple method to satisfy the first criterion and to provide pulse width proportional to e_1 is to use square hysteresis loop core material and proper circuitry. Square hysteresis loop core material is characterized by high remanence and exhibits both the property of memory and integration. The change in flux of such a core may be expressed by

$$\phi = K\int_0^{T/2} e_1 dt \text{ volt-seconds}$$

where ϕ is the total change in flux produced by a voltage e_1 applied to a winding of the core for a period of time $T/2$. If the core is always saturated at one knee of the hysteresis loop when the voltage e_1 is applied and the polarity of e_1 is chosen to produce a flux change toward the opposite knee, then the change in flux produced by e_1 will be proportional to the average value of e_1 for the time period $T/2$. The core will thus integrate the voltage e_1, and read in and store this volt-second integral as a flux level somewhere within the hysteresis loop. If, later, means are provided to resaturate the core or read out this stored flux with a constant voltage, then the time for read-out will be proportional to the amplitude of the voltage read-in, because the volt-seconds read-in will equal the volt-seconds read-out. In other words, the same value of the integral is obtained in the one case for a variable voltage and constant time and in the other case for a constant voltage and variable time. Fig. 1(A) and (B) illustrates the wave forms and circuitry used for this process. The supply voltage E_{ac} is a square wave of constant amplitude which establishes the read-in interval between time zero and time $T/2$ and read-out interval between time $T/2$

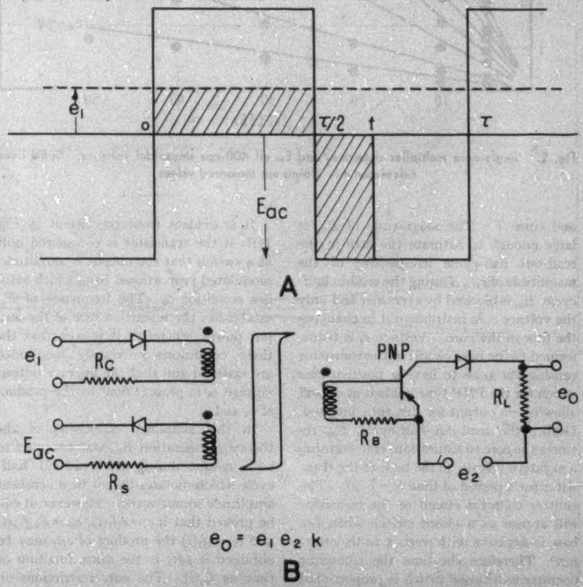


Fig. 1. A—Relationship of read-in voltage e_1 to read-out time $(t-T/2)$ for constant amplitude square wave E_{ac}. B—Simple single-core 1-quadrant multiplier

---

Paper 55-461, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 30, 1955; made available for printing April 20, 1955.

R. L. VAN ALLEN is with the Naval Research Laboratory, Washington, D. C.

The author wishes to acknowledge the contribution of D. H. Schaefer for the original conception of the basic idea of using a single core and transistor for obtaining pulse width and pulse amplitude modulation to perform computation.

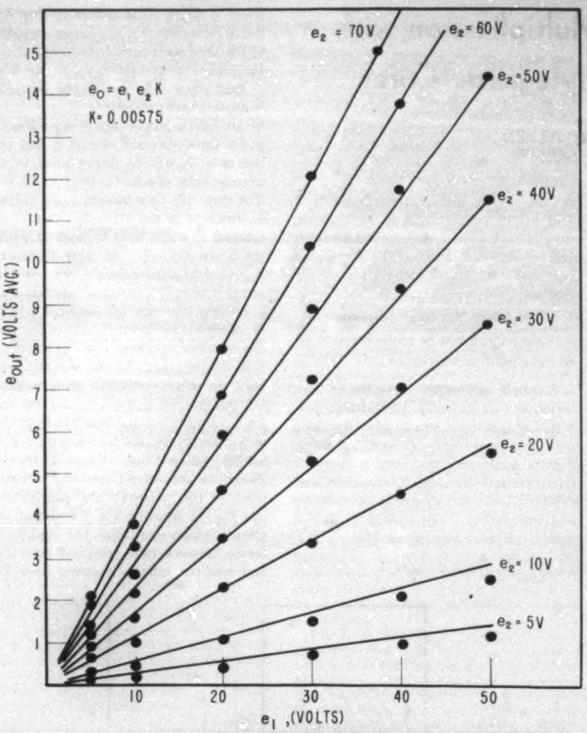


Fig. 2. Single-core multiplier data, $e_1 e_2$ and E_{ac} all 400-cps sinusoidal voltages. Solid lines calculated e_0. Points are measured values

and time T. The magnitude of E_{ac} is large enough to saturate the core in the read-out half-cycle irrespective of the magnitude of e_1. During the read-in half-cycle, E_{ac} is blocked by a rectifier and only the voltage e_1 is instrumental in changing the flux in the core. Voltage e_1 is transformed to the base circuit of the transistor causing the base to have a positive bias which for the PNP type holds it at cut-off allowing no output for this time interval. During the read-out half-cycle, E_{ac} restores the core to saturation, transforming a negative voltage to the base of the transistor for a period of time $(t-T/2)$. The emitter collector circuit of the transistor will appear as a closed switch when the base is negative with respect to the emitter.[3] Therefore the time the transistor operates as a closed switch is proportional to the input voltage e_1 which satisfies condition 1.

It is evident from the circuit in Fig. 1(B) if the transistor is considered only as a switch that the output is amplitude-modulated proportional to e_2, which satisfies condition 2. The frequency of E_{ac} establishes the repetition rate of the output pulses; therefore it is seen that the three conditions previously mentioned are satisfied and that the average output voltage e_0 is proportional to the product of e_1 and e_2.

In this simplified discussion of the theory of operation E_{ac} was assumed to be constant during the read-out half-cycle which dictates that it be a constant amplitude square wave. However, it can be proved that if $e_1 = e_1 f_1(t)$, $e_{ac} = E_{ac} f_{ac}(t)$ and $e_2 = e_2 f_2(t)$ the product of $e_1 e_2$ may be obtained if $f_2(t)$ is the same function of time as $f_{ac}(t)$. The only restrictions on the variable e_1 are that its volt-time average over the period of a half-cycle does not saturate the core, and within each half-cycle it does not reverse polarity. To realize the maximum range of values for the variable e_1 and to obtain half-cycle response for read-out, the magnitude of E_{ac} should be large enough to saturate the core in 1/2 cycle.

One-Quadrant Multiplier Using Sinusoidal Supplies

Fig. 1(B) bears a close similarity to a single core magnetic amplifier and the problems related to such devices are attendant here as well. Any lack of linearity of the magnetic amplifier circuit to produce a linear output will affect the linearity of the pulse width modulation as a function of the voltage input e_1. Another problem which affects the single-quadrant multiplier, but which is nonexistent in the 4-quadrant circuit, is the residual output when $e_1 = 0$. This is a result of uncontrolled reset in the core and is produced by rectifier back leakage as well as the effect of saturated inductance causing reset when the ratio of remanent flux to maximum flux is less than unity. This low-level inaccuracy dictates that the lower limit for e_1 must be about 5 volts, as shown in Fig. 2, which is a plot of calculated and experimental results. The solid lines are calculated values of e_0 using a multiplying constant of $K = 0.00575$, and the points indicated are experimental, measured values of e_0. These data were obtained with a breadboard circuit similar to that shown in Fig. 1(B). This is not purported to represent the optimum design for a single-quadrant multiplier but is given to show the circuit in its simplest form. Despite this, the maximum error is ±2% of full-scale values for $e_2 = 0$ to 70 volts and $e_1 = 5$ to 50 volts. The parameters for the circuit in Fig. 1(B) are as follows:

e_1, e_2, and e_{ac} supplied from a 400-cycle-per-second (cps) sinusoidal voltage source.
$R_c = 2500$ Ω, $R_s = 3500$ Ω, $R_B = 200$ Ω, $R_L = 500$ Ω.
Core *4179*, 1-mil tape Deltamax.
Read-in and read-out winding 1,000 turns each, base windings 200 turns.
All diodes are two *1N137-A* in series.
Transistor *2N43* PNP power junction.

Four-Quadrant Multiplication

A single-quadrant multiplier having been developed, it is not difficult to add components necessary to provide 4-quadrant multiplication. First, to provide for reverse polarity of e_1, another core must be added and, second, to provide for reverse polarity of e_2 a bidirectional, or a-c switch must be provided. A transistor

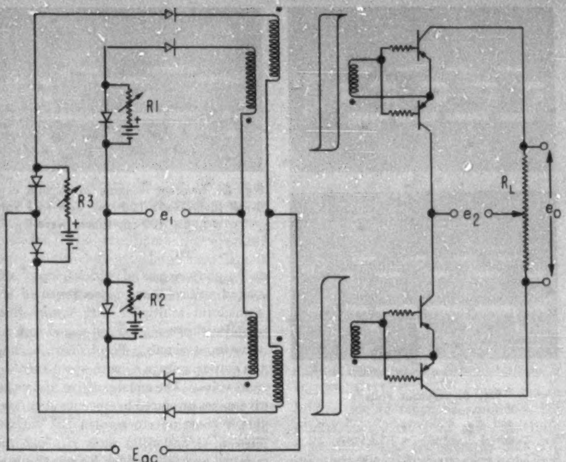
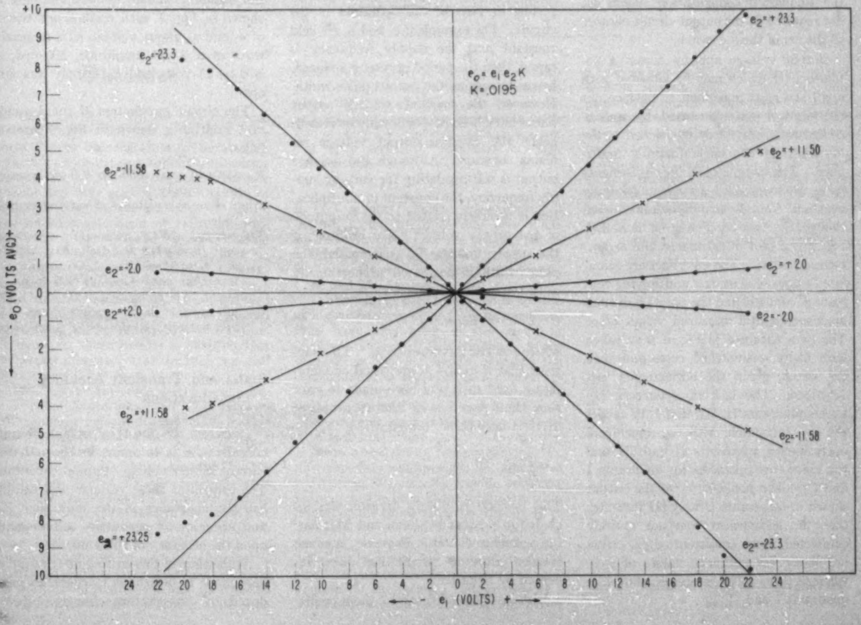


Fig. 3 (above). Four-quadrant multiplier for E_{ac} square-wave and e_1 and e_2 direct voltages

Fig. 4 (below). Four-quadrant multiplier data. Solid lines are calculated e_0. Points are measured values

bidirectional switch of the type reported by Bright[4] and Decker[5] serves admirably and exhibits the property of blocking or passing current in either direction. A complete 4-quadrant multiplier is shown in Fig. 3. The supply voltage E_{ac} is a square wave nominally 400-cps, which as previously mentioned restricts the wave form of e_2 to d-c or pulsating d-c if accuracy is to be maintained. Besides consisting of more components than the simple single-quadrant multiplier, there are several significant differences to be noted. The resistances R_c and R_s have been replaced with current-biased diodes to provide current-limiting in the back direction of the diodes. Substituting a current limiter for the supply resistance R_s prevents loading of E_{ac} when the cores have little or no resetting voltage e_1. Since on the read-out half cycle E_{ac} must supply magnetizing current plus leakage currents back through e_1, and base current to the appropriate transistor, the current limiter in the supply circuit must be set to pass several times magnetizing current. The control resistor R_3 for this current limiter[6] is made available for adjustment since it offers a limited control of the multiplying constant K in the expression $e_0 = e_1 e_2 K$. The current limiters



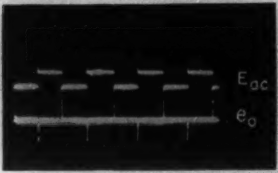


Fig. 5. Output voltage e_0 and E_{ac} 400-cps square wave for $e_1=0$ and $e_2=+23$ volts

in the input circuit for e_1 are kept separate to allow independent control (with R_1 and R_2) of the input impedance of each core presented to signal voltage e_1. When R_1 and R_2 are adjusted for slightly less than magnetizing current the effect is as though resistance has been added in the circuit on the read-in half-cycle. When unbalance exists between cores and transistors, it is possible to adjust the value of the constant K for two quadrants independent of the other two. For example, when R_1 is adjusted, the value of K is changed in quadrants II and III, or when R_2 is adjusted, the value of K is established for quadrants I and IV. This operation is analogous to changing the voltage gain of a magnetic amplifier by changing the value of the control resistance. For the 4-quadrant configuration, diodes are not required in the output circuit because of the use of the a-c switch.

Output voltage appears across a potentiometer which may be adjusted with $e_1=0$ and e_2 at maximum to balance out the residual output caused by uncontrolled reset previously mentioned in the discussion of the single-quadrant multiplier. This completes all the adjustments necessary to establish a single multiplying constant for all four quadrants even though the two cores are not matched. Fig. 4 is a plot of calculated and experimental results. The solid lines are calculated values of e_0 using a multiplying constant $K=0.0195$ and the points indicated are experimental measured values of e_0. The data obtained in Fig. 4 were taken with badly mismatched cores and then the circuit given the adjustments just described. The lack of symmetry between quadrants II, III, and I, IV reflect the core mismatch, with the upper core operating for quadrants II and III and the lower core operating for quadrants I and IV. The nonlinearity of the output shown in quadrants II and III indicates that the adjustment was not entirely completed, since consistent -2% errors are shown for midrange values of $-e_1$, whereas only $\pm 1\%$ errors are shown in quadrants I and IV.

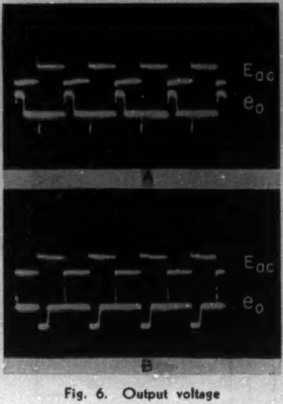


Fig. 6. Output voltage

A—e_0 and E_{ac} 400-cps square wave for $e_1=+8$ volts, $e_2=+23$ volts

B—Output voltage e_0 and E_{ac} 400-cps square wave for $e_1=-8$ volts; $e_2=+23$ volts

The constancy of the multiplier K was checked for different frequencies of the supply voltage E_{ac} and found to vary approximately 1% over a 2-to-1 frequency range of 200 to 400 cps. This is not surprising in view of the operation of the circuit. For example if e_1 and e_2 are held constant and the supply frequency is raised, then the period of reset is reduced, hence decreasing the output pulse width. However, the frequency of the output pulses have been proportionally increased, hence the average output voltage remains constant. Although the average output is self-regulating for varying supply frequency, the constant of multiplication is inversely related to the amplitude of the supply voltage. For example, if the supply voltage E_{ac} is increased, the pulse width is decreased proportionally as shown in the relationship

$$(t-T/2)=\phi/E_{ac}$$

where ϕ is the flux reset by e_1. This suggests that if e_1 and e_2 are equal or maintain a fixed ratio and both made to vary as a third function e_n, the operation of division may be performed where

$$e_0=\frac{e_a}{E_{ac}}K$$

This method to obtain division was included in a paper by Finzi and Mathias[2] for a similar circuit. However, a much simpler dividing circuit has been invented by Schaefer.[6]

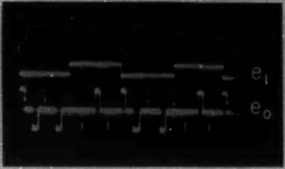


Fig. 7. Transient response of e_0 with $e_1=$ 5-volt square-wave 100-cps; $e_2=+23$ volts d-c, E_{ac} 400-cps square wave

Figs. 5, 6(A), 6(B), and 7 show oscilloscope photographs of various input and output voltage wave forms taken of the 4-quadrant multiplier with square-wave supply voltage E_{ac}, and d-c and square-wave input signals. Fig. 5 shows E_{ac} and the output voltage e_0, with $e_1=0$ and $e_2=\pm 23$ volts. The equal positive and negative spikes produced by uncontrolled reset flux average out to zero in the output. Figs. 6(A) and 6(B) show E_{ac} and the polarity and pulse width for positive and negative d-c input voltages; Fig. 6(A) shows $e_1=+8$ volts, and $e_2=+23$ volts; and Fig. 6(B) shows $e_1=-8$ volts and e_2+23 volts.

To obtain a measure of the dynamic accuracy the speed of response of the circuit was checked with a square-wave input signal. Inherent 1-cycle response is shown in Fig. 7 with oscilloscope traces of e_1 and e_0, input voltage e_1 is a square wave of a 5-volt amplitude, 100 cps; e_2 is d-c+23 volts and E_{ac} supply was 400 cps.

The circuit parameters of the 4-quadrant multiplier shown in Fig. 3 are as follows:

$E_{ac}=24$-volt peak, 200- to 400-cps square-wave supply.
e_1 and e_2 are d-c voltages of varying magnitude.
Transistors, *2N43* germanium junction PNP, cores *4179 D-1* Deltamax, supply and e_1 input windings 1,000 turns each, transistor base windings 200 turns.
Transistor base resistors 470 Ω, $R_L=1{,}000$ Ω, current limiter supply three 45-volt batteries, diodes *1N93* germanium junction.

Static and Transient Accuracy Considerations

Probably the most important circuit consideration is to maintain E_{ac} and the current limiter voltage supplies constant. The control of these voltages will establish the constancy of the multiplier K, and accuracy of operation will depend upon the accuracy of this control.

At the time of this writing no tests have been made to determine inaccuracies produced by temperature changes. However, the diodes would be expected to introduce the largest errors, especially those in the current limiter circuits for e_1, since they are adjusted to block magnetizing current partially in the e_1 control circuit. Since these diodes are shunted by resistors R_1 and R_2, an excellent opportunity is afforded to obtain temperature compensation. Temperature effects on the cores are expected to be second-order effects in the temperature range dictated by the use of germanium components. The transistors are operated only as switches with large base signals and again only second-order effects would be expected from similar temperature variations.

The source impedances of the signal voltages are significant in that e_1 is required to supply magnetizing current and E_{ac} must supply at least twice magnetizing current. For the circuit shown in Fig. 3, E_{ac} delivered a maximum of 13 milliamperes from a 1,000-Ω source and e_1 delivered 5 milliamperes from a 200-Ω source. There are no special limitations on the source of e_2 as long as its impedance is substantially less than the leakage impedance of the transistor a-c switch when blocking. The leakage of these switches was checked with zero base driving voltage and found to be about 600,000Ω.

The requirement that e_2 must have the same function of time as E_{ac} has the effect of limiting the bandwidth of the input voltage e_2. For example, when a square-wave supply E_{ac} is used, e_2 is restricted to direct current, square wave, or a slow time-varying direct current (low-frequency alternating current) with accuracy decreasing with increasing frequency of e_2. However, if e_2 is supplied by a modulator having a carrier with a wave form of the same time function as E_{ac} and amplitude modulated proportionally to a signal $e_2{}'$ the bandwidth for the signal $e_2{}'$ may be greatly increased. However, as in any carrier-modulated device the bandwidth must be always less than the frequency of the carrier. Considerations of bandwidth are important in many computer problems. But it should be remembered that the output of this device is in the form of pulses and, as in the case of magnetic amplifiers, has similar limitations. In this, the transient response is the same as that of a good half-wave magnetic amplifier with a minimum response of 1/2 cycle and a maximum of 1 cycle, as shown in Fig. 7.

Summary

It has been shown that a 4-quadrant analogue multiplier can be built with the use of conventional magnetic-amplifier components plus four transistors. The performance of the first breadboard model gave accuracy of ±1% over a range of 0 to 16 volts for each input variable, with good prospects of improvement with core matching.

The output voltage of the device is in the form of d-c pulses with reversible polarity depending upon the sign of the input voltages, each pulse being width-modulated proportional to one of the input voltages and amplitude-modulated proportional to the other.

References

1. THE SINGLE-CORE MAGNETIC AMPLIFIERS, AS A COMPUTER ELEMENT, **R. A. Ramey.** *AIEE Transactions*, vol. 71, pt. 1, 1952, (Jan. 1953 section), pp. 442–46.

2. AN APPLICATION OF MAGNETIC AMPLIFIER CIRCUITS TO PERFORM MULTIPLICATION AND OTHER ANALYTICAL OPERATIONS, **L. A. Finzi, R. A. Mathias.** *AIEE Transactions*, vol. 72, pt. I, Sept. 1953, pp. 455–61.

3. JUNCTION TRANSISTORS USED AS SWITCHES, **R. L. Bright.** *AIEE Transactions*, vol. 74, pt. I, Mar. 1955, pp. 111–21.

4. A SWITCHING TRANSISTOR D-C TO A-C CONVERTER HAVING AN OUTPUT FREQUENCY PROPORTIONAL TO THE D-C INPUT VOLTAGE, **G. H. Royer.** *AIEE Transactions*, vol. 74, pt. 1, July 1955, pp. 322–26.

5. TRANSISTOR DEMODULATOR FOR HIGH-PERFORMANCE MAGNETIC AMPLIFIERS IN A-C SERVO APPLICATIONS, **R. O. Decker.** *AIEE Transactions*, vol. 74, pt. I, Mar. 1955, pp. 121–23.

6. A STATIC MAGNETIC TRANSISTOR ANALOGUE DIVIDER, **D. H. Schaefer.** *AIEE Transactions*, vol. 74, pt. I, 1955 (*Paper No. 55–451*).

◆

Discussion

K. Chen and **R. O. Decker** (Westinghouse Electric Corporation, Pittsburgh, Pa.): The author is to be congratulated for making effective use of the new switching transistor techniques. The switching transistor should find wide application in the field of analogue computation and control. Since the initial development of switching transistor circuitry, we have been working on analogue computing techniques utilizing the unique properties of switching trantransistors. To date, many schemes have been developed, including 2-quadrant and 4-quadrant multipliers, which use different circuitry from that described by the author. A paper on these circuits is being written by use for presentation at the 1955 National Electronics Conference.

One major difference between our circuitry and that of the author is the use of an all-transistor pulse-width modulator. This was developed after experimenting with a circuit similar to that of Fig. 3. The linearity of the multiplier described by the author is limited by the magnetic core properties and we have experienced difficulties with the change in core characteristics with temperature. High-frequency operation also seems to be difficult with the magnetic pulse-width modulator because of increased exciting current and deterioration of sharp firing angle at the high frequencies.

Another drawback of the circuits contained in this paper is the lack of a blocking voltage for the switching transistors when the cores saturate. It is our experience that switching transistors with zero or very low blocking voltages are not idal open switches, especially at high temperatures.

•

R. L. Van Allen: Further development of the 4-quadrant multiplier has been made since the submission of the original paper. A brief summary of this work will serve, in part, to answer the points raised by Mr. Chen and Mr. Decker.

The circuit has been redesigned to operate at higher frequencies and several modifications have been made in the circuit to improve linearity. This modified circuit is shown in Fig. 8. Besides changing the core material to Hy Mu 80, the most significant change is to remove the current limiters associated with R_1 and R_2 of Fig. 3. These current limiters have been replaced with switching transistors. A transformer has been added with a primary winding W_1, and secondary windings W_2, W_3, and W_4.

Windings W_2 and W_3 provide switching currents in the base of the additional transistors of the proper polarity effectively to decouple the read-in circuit containing e_1 and the read-out circuit containing E_{ac}. Winding W_4 has a few turns to provide a small voltage in the read-in circuit to compensate for the forward voltage drop of a *2N44* transistor and a *1N93* rectifier during a read-in half-cycle. The polarity of this winding is such that it effectively compensates both cores in this respect. If high temperatures are anticipated it is desirable to provide a positive cutoff voltage to the base of the transistors in the output circuit; shown as a $1^1/_2$-volt bias in Fig. 8. However, a simpler arrangement would be to omit the battery and simply by-pass the base resistors with capacitors. Since leakage of the four transistors in the bridge-type output circuit will not vary uniformly

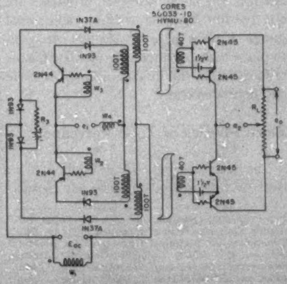


Fig. 8. Modified circuit for 5-quadrant multiplier

with temperature this voltage is desirable as mentioned by Mr. Chen and Mr. Decker.

The modified circuit of Fig. 8 has been operated using a 5-kc square-wave supply and with e_1 source impedance of 1,000 ohms. The over-all accuracy was slightly better than ±1% at room temperature. Further tests were made to determine the effects of temperature on the cores by subjecting them to temperature changes while maintaining the rest of the circuit at constant temperature. When the temperature of the cores was cycled from 0 F to 180 F over-all accuracy did not vary more than ±2%. This is not too surprising since despite individual core changes the theory of operation in this multiplier is that, if n volt-seconds are read into a core, then n volt-seconds are read out of the core.

When the entire circuits, including the cores, was subjected to temperature from 0 F to 140 F, the error was as much as 4% and not uniform in the four respective quadrants, indicating that either the transistors and/or the diodes were contributing error. At the present time complete information has not been obtained on the temperature characteristics of the individual components of the circuit excluding the cores.

In regard to high-frequency operation and the associated increase in magnetizing current mentioned by Mr. Chen and Mr. Decker, it has been found that to maintain good linearity, the source impedance of the read-in circuit must be kept low. For example, at 5 kc in order to maintain 1% accuracy it was not practical to have the read-in source impedance exceed 1,000 ohms.

Notes on the Use of Screen-to-Plate Transconductance in Multigrid Tube Circuit Design

K. A. PULLEN, JR.
MEMBER AIEE

THE conductance contour technique[1] of circuit design provides a direct and rapid means of co-ordinated static and dynamic design of electronic circuits. It yields results in a form which enables the designer to tell by inspection whether the design will comply with design requirements or is marginal or unsuitable. Design rules based on the conductance technique enable the user to test alternate tubes for the tubes having unsatisfactory characteristics. However, the rules do not make it possible to reconsider the problem and directly select two or three tubes offering the best opportunity for achieving the desired design.

It might be supposed that the division of tubes into low-, medium-, and high-power groups would provide the separation required for adequate design. Unfortunately, with such a tube-type classification, the major selection problems are not readily handled without additional design criteria or aids. Data on screen-to-plate transconductance offer one means of facilitating the final choices for tetrode and pentode tubes.

Basic Physics of Tube Perveance

The term "perveance" as applied to an electron tube is used qualitatively to separate tubes by their basic diode connection current-voltage characteristics. Tubes capable of developing large cathode currents within their electrode power dissipation ratings are classified as high-perveance tubes, the rest as medium- or low-perveance tubes. Actually, classification by perveance accomplishes little not already accomplished by the titles high-, medium-, and low-power tubes. The classification of both systems in general tends to place high-power and high-perveance tubes together and similarly with the remaining classifications.

The problem of the selection can best be attacked by learning the basic significance of perveance of a tube in relation to its power-handling ability. Then the tube parameter or parameters best indicating the properties which need to be represented can be selected.

What is the basic difference between two tubes having major parameters identical with the exception of one single basic parameter.[2] Several sets of tubes might be set up for comparison to aid in isolation of the appropriate parameter. The *6SN7*, the *12AU7*, the *5687*, and the *12BH7* tubes might be selected as one comparison set. These double triode tubes have approximately the same amplification factor. Otherwise they have a wide range of power-handling abilities. Consequently, the tube amplification factor is not of prime importance to power-handling ability. On the other hand, comparison of the *12BZ7* and the *12BH7* tubes, Figs. 1 and 2, shows the former to have low power-handling ability, and the latter a higher power-handling ability.

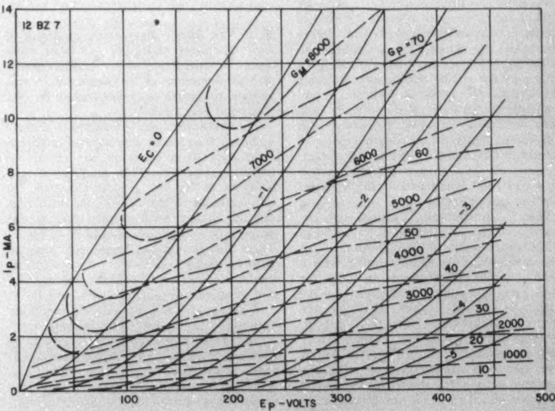


Fig. 1. 12BZ7 plate characteristic curves

Paper 55-500, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 2, 1954; made available for printing May 11, 1955.

K. A. PULLEN, JR., is with the Ballistic Research Laboratories, Aberdeen Proving Ground, Md.

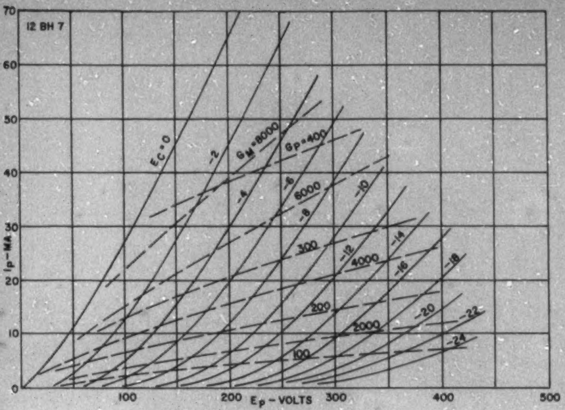


Fig. 2. 12BH7 plate characteristic curves

The difference, as is now generally realized is primarily one of plate conductance, since the transconductances of the two tubes are largely comparable.

It might be asked just how the plate conductance difference is able to produce such a large difference between two tubes which have nearly equivalent transconductance characteristics. The answer to this question is of fundamental importance to the understanding of the ability of a tube amplifier to handle high-power signals.

If an experimental triode tube were to be built on which the transconductance were fixed and the plate conductance variable, how would the characteristics of the tube change as the plate conductance was varied? Examination of the typical plate characteristic curves of the *12BH7* and the *12BZ7* tubes shows that the effect of increase of plate conductance is to cause a more rapid rise of plate current with plate voltage (higher perveance). As the plate conductance is increased, more and still more bias contours appear near zero plate current. The tube requires more bias to cut off the plate current. As either the plate supply voltage or the plate conductance is increased, the total change of plate current available and the change of bias available in the tube increase.

In tetrode and pentode tubes a similar situation exists. In multigrid tubes, the plate voltage has relatively little effect on the plate current. Variation of the screen voltage applied to the tetrode or pentode tube shows that the parameter in pentodes corresponding to triode plate conductance is screen-to-plate transconductance. As in the case of triode tubes, tetrodes or pentodes which control large amounts of current must either have a high screen-to-plate transconductance, or a high screen (and even higher plate) voltage. The correctness of this conclusion can be readily verified by examination of the screen characteristic curves of the *12BY7* and the *6BQ6* tubes; see Figs. 3 and 4. Estimating the screen-to-plate transconductance along the zero bias line for these tubes gives approximate values of up to 500 micromhos for the *12BY7* tube and up to 2,500 micromhos for the *6BQ6* tube. Consequently, the current-change capacity of the *6BQ6* tube is five times that of the *12BY7* tube. The bias range is correspondingly altered.

The next question to be answered is how much of the current-change capacity can be used successfully? Here the question of plate and/or screen dissipation becomes important. The factor of importance with the *12BY7* and *6BQ6* tubes is the plate dissipation.

Assume that the design for the use of the *12BY7* tube has shown that the *12BY7* design is marginal in operation. If the *6BQ6* tube is substituted for the *12BY7*, the change in plate-dissipation capabilities would permit up to 70-percent increase of static plate current. The tube, inherently, however, is capable of five times the current change of the *12BY7*. Obviously, the circuit should be redesigned to provide the maximum current-change capabilities. Because the power-handling ability of the *6BQ6* is only 70 per cent greater than that of the *12BY7*, the full capabilities for apparent current change cannot be realized. The possible reduction of screen and plate voltage may make available up to $2^1/_2$ times the current change available in the *12BY7* tube.

A design can be altered by selection of a tube having the same screen-to-plate transconductance but higher plate dissipation. Such a change might be the substitution of a *6CL6* tube for a *12BY7*. The advantage gained is frequently of rather small value unless the actual increase in dissipation available is of the order of two or three times.

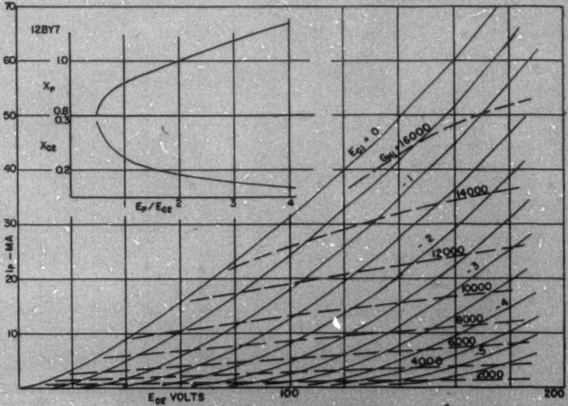


Fig. 3. 12BY7 screen characteristic curves

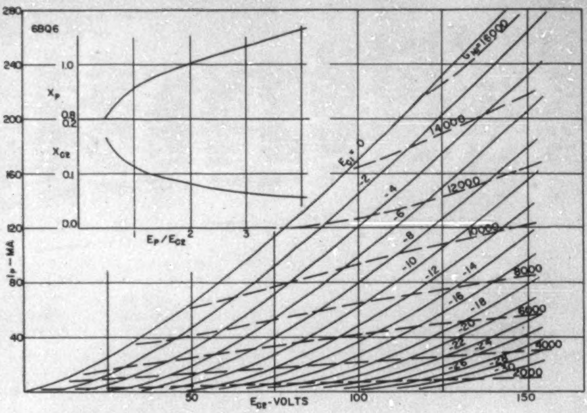


Fig. 4. 6BQ6 screen characteristic curves

Using Screen-to-Plate Transconductance in Design

The value of screen-to-plate transconductance required of a pentode depends entirely on the signal characteristics required of a circuit. Any given type pentode tube can be made to accomplish a moderate range of operations. Where design for reliability is required, a series of pentode tubes should be established which have screen-to-plate transconductance geometric ratios of 2 to 4 with respect to each other, and a gradually increasing power-handling ability as the screen-to-plate transconductance is increased. Then the lowest screen-to-plate transconductance tube of the series capable of meeting distortion and power-handling limitations with sufficiently small dissipation requirements is selected to provide reliable and economical operation.

The trend to compact, lightweight equipment for both military and commercial applications has made an understanding of selection of tubes on the basis of screen-to-plate transconductance more than normally important. Stringency of cooling problems makes it necessary for the designer to keep tube and component dissipations as low as possible. Keeping dissipation low requires the use of adequately high screen-to-plate transconductance tubes in the higher power stages of electronic equipment.

A tabulation of typical values of maximum screen-to-plate transconductance on some commercial tetrode and pentode tubes probably would be useful to many engineers. This is contained in Table I, with the tubes listed in order of ascending capacity. The variations of control-grid transconductance indicate the ease of drive of the tube, and the screen-to-plate transconductance, the power-handling ability. It is interesting to note that although tubes 5, 6, and 8 have to the author's knowledge been used for cathode-ray oscillograph final deflection amplifiers, their use has forced the development of high vertical-sensitivity cathode-ray tubes. Tubes 9, 10, and 11, which have been little used as final deflection amplifier tubes, largely eliminate the need for special high deflection-sensitivity cathode-ray tubes.

Table I. Pentode Power Design Parameters

Tube No.	Tube Type	Typical G_{M1}, Micromhos	Maximum G_{M2}, Micromhos	Plate Power Dissipation, Watts
1	*6BH6*	4,500	105	3.0
2	*6CB6*	7,000	200	2.0
3	*6AK5*	6,000	267	1.7
4	*6AH6*	12,000	267	3.0
5	*12BY7*	14,000	500	6.0
6	*6AG7*	11,000	526	9.0
7	*6V6*	5,000	588	12.0
8	*6L6*	7,000	800	19.0
9	*6216*	14,000	2200	10.0
10	*6Y6*	12,000	2200	12.5
11	*6BQ6*	14,000	2500	10.0

Screen-to-Plate Transconductance in Push-Pull Amplifiers

If a quantity of tubes of a given *RMA* type is tested, it is found that, whereas a large proportion of the tubes will test reasonably close to one another on screen-to-plate transconductance, a significant number of tubes, while testing normal on control grid transconductance, will test from 10 to 50 per cent either high or low on screen-to-plate transconductance. A typical case of this sort is shown in Fig. 5. Tube no. 3 represents the average or bogie tube. Number 9 represents one having an abnormally high screen-to-plate transconductance (over 30 per cent high). The data on no. 10 show the effect of low emission on an otherwise normal tube. The results of attempting to use tubes 3 and 9 in push-pull, particularly in a d-c amplifier, can be imagined. Control of the screen-to-plate transconductance is consequently important for tubes used in d-c amplifiers.

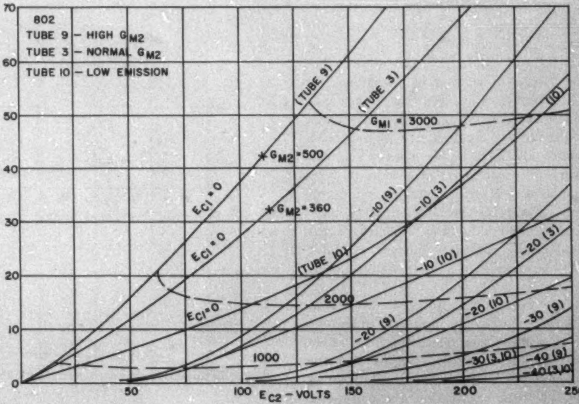


Fig. 5. 802 comparative screen characteristic curves

Some notes on the testing of tubes 3 and 9 further indicate the importance of screen-to-plate transconductance. If both the tubes are tested with a transconductance-type tester, both tubes will test to be good. Testing with an emission-type tester will test both tubes as good, with tube 9 testing as even better than tube 3. The significance of the difference between these two tubes can be detected only by a more elaborate test such as plotting two or three of the typical screen contours, as shown in Fig. 5. The typical result of an increase of screen-to-plate transconductance in the bias contours for tube 9 is the abnormal swinging upward of the individual bias contours as the screen voltage rises with additional low-current area bias contours introduced.

Conclusions

The foregoing discussion of uses of the screen-to-plate transconductance parameter and its application is by no means complete. It indicates some of the more directly useful applications of the concept. Experience indicates that in most applications only an approximate comparative value of the parameter need be used in combination with plate-dissipation data to select tubes for given applications. When more detailed data are required, they may be estimated from the slopes of the bias contours on the screen characteristic data plots.

An additional use of the screen-to-plate transconductance parameters is in the matching of tubes for push-pull amplifiers. Failure to maintain proper matching can easily result in improper operation of such a device as an oscillograph d-c amplifier. Some special balancing techniques using simple and readily available test equipment need to be developed for balance testing. Emission testing may give a clue to trouble in balance but may, because of self-bias arrangement, prove inadequate.

The screen-to-plate transconductance parameter, like its near relative, the triode-plate conductance, is of great circuit-design importance. Its significance, like that of the plate conductance, is lost in the establishment of the screen grid μ factor. Like the triode μ factor, the screen grid μ factor reduces to a ratio two of the most important parameters controlling utilization of pentode tubes. Consequently, this paper has been prepared to clarify the usefulness and importance of screen-to-plate transconductance as a fundamental parameter in circuit design.

References

1. Instrument Design Using G Curves, K. A. Pullen. *The Instrument Maker*, vol. 17, no. 1, 1949, pp. 4–5.

2. Amplifier Design Using G Curves, K. A. Pullen. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 37, no. 2, 1949, p. 210.

3. Using G Curves in Tube Circuit Design, parts 1 and 2, K. A. Pullen. *Tele-Tech*, New York, N. Y., vol. 8, nos. 7 and 8, 1949, pp. 34–6, and pp. 33–5 and 59.

4. The Use of Conductance Curves for Pentode Circuit Design, K. A. Pullen. *Ibid.*, vol. 9, no. 11, 1950, pp. 38–40, 45.

5. Notes on UHF Oscillator Design, K. A. Pullen. *Ibid.*, vol. 12, no. 2, 1952, pp. 80–1, 142–5.

6. Conductance Curves Speed Triode R-C Amplifier Design, K. A. Pullen. *Ibid.*, no. 5, 1953, pp. 80–1, 123, 126–8.

7. Conductance Curves Speed Pentode R-C Amplifier Design, K. A. Pullen. *Ibid.*, no. 7, 1953, pp. 44–6, 86, 92.

8. G Curves and Impedance Amplifiers, K. A. Pullen. *Ibid.*, no. 9, 1953, pp. 71, 142–7.

9. Tube Circuit Design Using the G Curve Technique, K. A. Pullen. *Memorandum Report No. 489*, Ballistic Research Laboratories, Aberdeen Proving Ground, Md., 1949.

10. The Use of Conductance Curves for Pentode Circuit Design, A. H. Hodge, K. A. Pullen. *Memorandum Report No. 499, Ibid.*, 1950.

11. Improved Techniques for Tube Circuit Design, K. A. Pullen. *Radio and Television News*, Chicago, Ill., vol. 46, no. 1, 1951, pp. 10–12, 27–8.

12. Using Conductance Curves in Electronic Circuit Design, K. A. Pullen. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 6, 1950, pp. 112–20.

13. The Conductance Curve Design Book, K. A. Pullen. Kann-ellert Electronics, Baltimore, Md., 1953.

◆

Discussion

D. C. Jackson, Jr. (Ballistic Institute, Ordnance Corps, Aberdeen Proving Ground, Md.): The data on screen-to-plate transconductance discussed in this paper appear very interesting. I am particularly interested in the statement that the new transconductance parameter is of use in design for reliable circuit operation. This is a point that I believe should be more fully treated than was done in the paper and I should like the author to discuss it in greater detail.

K. A. Pullen, Jr: The point raised by Prof. Jackson is very important. I regret that lack of space prevented a more full discussion. The new transconductance parameter enables one to select substitute tubes which often permit reduction of power dissipation and provide improved reliability.

As an example, consider the problem of design of a circuit using a pentode tube for series regulation service in a regulated power supply. The pentode tube shall pass a current between 20 and eighty milliamperes.

Three methods are presently available to select the tube capable of providing the required current:

1. Use of cathode current capacity data and typical curves as given in tube handbooks.
2. Use of data on maximum plate current and screen current at rated dissipation in conjunction with typical tube data.
3. Use of screen-to-plate transconductance G_{M2} in conjunction with screen characteristic curves.

Use of the first method enables one to find the tubes capable of handling the current, but only by tedious searching. After the tubes capable of meeting performance requirements have been selected, then it is necessary to select the preferred tube of the group and to establish the required screen and plate voltages for the tube. With the tube characteristic data normally available, cut-and-try methods must be used for adjustment of the screen voltage and the plate voltage.

Use of the data on maximum plate current and screen current within tube ratings suffers from rather similar limitations. At its present state of development, the same difficulties exist in selection of proper operating voltages. Whereas the method does appear to be capable of some development which may make it useful, at present it does not provide a direct solution.

In contrast to the foregoing methods, the use of the screen-to-plate transconductance parameter gives a consistent and co-ordinated approach to the problem. In making preparation to select the preferred tube, first the maximum required current is taken and divided by the nominal G_{M2} for the tube under consideration. This gives the mini-

Table II. Results of Calculations

Tube	$P_{p\max}$, Watts	G_{M2}	I_{bsp}/G_{M2}	$1.5I_{bsp}/G_{M2}$	E_{c2}	E_{bb}	P_p, Watts
6L6	19	0.0008	100	150	150	150	12
6216	10	0.0022	37	56	56	56	4.5
6Y6	12.5	0.0022	37	56	56	56	4.5
6BQ6	10	0.0025	32	48	48	48	4.0

Note. I_{bsp} is the peak nominal plate current, or 80 milliamperes. As can be seen, the *6L6* is operating at $^{2}/_{3}$ of full rating, whereas the remaining tubes are all operating at less than half of full rating.

Table III. Required Screen Voltages

Tube	E_{c2} Calculated, Volts	E_{c2} From Chart, Volts
6L6	150	150
6216	56	70
6Y6	56	60
6Y6	56	60
6BQ6	48	63

mum screen voltage required in a tube having bias contours whose slope as a function of screen voltage is constant. Since the bias contour is actually curved, it is necessary to multiply the minimum screen voltage obtained by 1.5 to allow for the curvature. The plate voltage is then chosen to be approximately equal to the screen voltage (for convenience only). Tabulations of the results of calculations on tubes 8 through 11 in Table I are shown in Table II.

A check of the typical screen characteristic curve sheets of the four tubes (the curve data on the *6BQ6* are included in the paper) shows that the screen voltage required to produce 80 milliamperes of current in the tubes are as given in Table III. The calculated values for the *6216* and the *6BQ6* are somewhat low because of the high curvature of the bias contours at low voltage. In spite of the curvature, it is interesting to note that the predicted voltage is close to the required voltage. Consequently, selection of the correct tube can be made quickly. As has already been noted, however, use of the first two methods does not give the user even an inkling as to what the correct screen and plate voltages might be, let alone permit the user to find the tube which will meet design requirements with minimum tube dissipation and minimum equipment heating.

Although the plate voltage has been selected to be nominally equal to the screen voltage, any plate voltage between about three-quarters and twice the screen voltage could be used. The only difficulty with the higher values is the necessity to be able to handle more power than is necessary.

Since one of the important problems of circuit reliability is keeping over-all dissipations as low as is consistent with design requirements and with operation within the controlled areas for the tubes, the use of screen-to-plate transconductance should prove valuable to the design engineer.

A Static Magnetic Transistor Analogue Divider

D. H. SCHAEFER
ASSOCIATE MEMBER AIEE

RECTANGULAR loop magnetic materials and transistors used as switches have been combined to produce a simple, high-speed analogue divider. This device is capable of producing an output whose average voltage is the quotient, with the correct sign, of two input voltages. A preliminary model of this divider gives accuracies of ±2.5 per cent of full scale readings for a quotient range of 100 to 1 and input variations of both numerator and denominator of better than 10 to 1.

Theory of Operation

The divider produces an average output voltage that is proportional to the quotient of two input voltages by having the numerator determine the magnitude of the output voltage while the denominator determines the amount of time this voltage is allowed to appear across the output terminals.

The element that directly determines the amount of time the output will be present is a square-loop magnetic core. To drive any magnetic core from saturation in one direction to saturation in the other direction requires a given number of volt-seconds. In the case of square-loop materials, there is a very sudden and sharp point of saturation where the voltage across the core winding will drop to zero.

The time of saturation of a magnetic core whose primary is being fed by a square-wave voltage generator, Fig. 1, will now be determined. The rectangle between the windings of Fig. 1 indicates that the core is of square-loop material. Assuming e_d at least large enough to drive the core from saturation to saturation on each of its half-cycles, and also assuming the drop across the resistance in the primary circuit to be negligible before saturation, the following equation is true

$$\int_0^{t_s} e_d dt = K \tag{1}$$

where

time is measured from the beginning of a half-cycle

t_s = the time of saturation

K = the number of volt-seconds required to drive the core from saturation in one direction to saturation in the opposite direction

As e_d is a square wave, during any half-cycle it is a constant and can be brought outside of the integral sign. Equation 1 can therefore be simplified to

$$e_d t_s = K \tag{2}$$

or

$$t_s = \frac{K}{e_d} \tag{3}$$

The time taken to drive a core from saturation to saturation is therefore inversely proportional to the voltage applied to it or, more plainly, the larger the voltage across a core the sooner it will saturate.

If a switch in a circuit, such as Fig. 2, could be so controlled so that it would be closed only when the core of Fig. 1 were in an unsaturated condition, then the average voltage across the resistor of Fig. 2 would be proportional to e_n/e_d. Transistors make excellent switches and their properties when used as switching elements is thoroughly covered in reference 1. The basic rule to keep in mind for PNP type of transistors is that the collector-to-emitter impedance of a transistor becomes very great if, and only if, both the collector and emitter voltages are equal to or more negative than the base voltage. If the base, on the other hand, is on the order of a few tenths of a volt negative with respect to either the collector or the emitter, the emitter-collector impedance can be as low as 1 ohm.

Fig. 3 is the basic version of the dividing circuit. In the circuit, the transistor is being used as a switch. If the voltage on the core windings is positive, the transistor will be in a conducting state and the numerator voltage will appear across R_L. If the core windings have a negative or zero voltage, the transistor

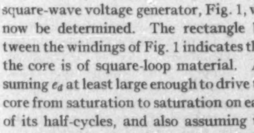


Fig. 1 (left). Denominator circuit of divider

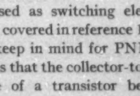


Fig. 2 (right). Equivalent of numerator circuit

Paper 55-451, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 30, 1955; made available for printing April 18, 1955.

D. H. SCHAEFER is with the Naval Research Laboratory, Washington, D. C.

The author gratefully acknowledges the suggestions and advice of Roland Van Allen and the assistance of Paul Burris who built and obtained the measurements upon the experimental circuit.

Vol.
Pt.
195

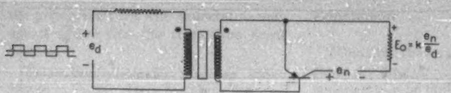


Fig. 3 (left). The single - quadrant dividing circuit

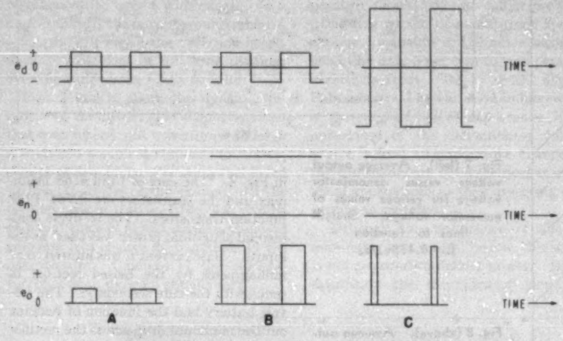


Fig. 4. Typical inputs and outputs for circuit of Fig. 3

will be in a nonconducting state and no voltage will appear across R_L. Therefore, the voltage e_n will appear across R_L whenever the denominator voltage e_d is positive and the core is not saturated. Three typical inputs and outputs for a circuit such as Fig. 3 are shown in Fig. 4. It should be noted that the ratio of e_n to e_d is equal in Figs. 4(A) and (C) and that the average value of e_0 is equal, as it should be, in these two figures.

To obtain a 4-quadrant divider, the circuitry of Fig. 5 is utilized. This circuit was originated by D. G. Scorgie and utilizes the 2-transistor a-c switch described by Bright and Decker.[1,2] In Fig. 5 both the numerator and denominator voltages are in the form of square waves. The two transistors are necessary to insure that e_n can be blocked on both its positive and negative half-cycles. If e_n and e_d are in phase, the transistors will allow e_n to appear across R_L only during its positive half-cycles; therefore e_0 will be positive. If e_n and e_d are 180 degrees out of phase, the transistors will allow e_n to appear across R_L only during negative half-cycles. If e_d and e_n are, e.g., outputs from two ring modulators with the same a-c supply, it can be seen that the divider of Fig. 5 will give the quotient, with the proper sign, of the d-c inputs to the ring modulators.

SIGNAL WAVE SHAPES

The numerator and denominator voltages of Fig. 5 do not have to be square waves. They can be sinusoids or any other type of periodic alternating functions, the only restriction being that e_n and e_d must have the same functional form. The proof of this for the steady-state situation is as follows

Given:

$e_d = E_d f(t)$

$e_n = E_n f(t)$

where

E_d, E_n = constants as the steady-state situation is being considered

$f(t)$ = a periodic alternating function of period greater than t_s

To prove: the average value of the output is proportional to

$$\frac{E_n}{E_d}$$

Proof:

$$\int_0 E_d f(t)dt = K \qquad (4)$$

where

t_s = the time of saturation

K = the number of volt-seconds required to go from saturation to saturation

Considering E_d a constant

$$E_d \int_0^{t_s} f(t)dt = K \qquad (5$$

or

$$\int_0^{t_s} f(t)dt = \frac{K}{E_d} \qquad (6)$$

If $\bar{E}_0$ is the average value of the output per cycle

$$\bar{E}_0 = \int_0^{t_s} E_n f(t)dt \qquad (7)$$

$$= E_n \int_0^{t_s} f(t)dt \qquad (8)$$

$$= K\frac{E_n}{E_d} \text{ by equation 6, Q.E.D.} \qquad (9)$$

In the foregoing derivation it was considered that E_n and E_d were constants. It can be seen, however, that if either of these coefficients underwent a step change in value the output would assume its new value within 1 cycle of $f(t)$, the exact amount of time depending upon when the step change occurs in the cycle. If E_d and E_n are continuously varying with time, the output is the quotient of the averages of the inputs during the time the core is unsaturated. In essence, the divider samples the inputs once a cycle, the duration of time over which the sample is taken being dependent upon the magnitude of the denominator.

Practical Considerations

In deriving equation 3 it was assumed that the drop across the resistance in the primary circuit was negligible. In an actual circuit this drop is present and assumes importance as the denominator signals get smaller. To see this, it should be noted that the core sees the voltage $e_d - iR$, rather than simply e_d, where R = resistance in the denominator circuit and i = current flowing in this circuit. Therefore the output will be proportional to $e_n/(e_d - iR)$. The iR

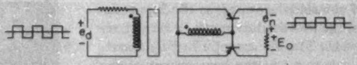


Fig. 5 (right). The 4-quadrant dividing circuit

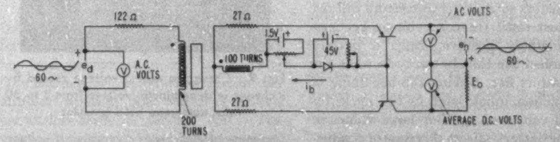
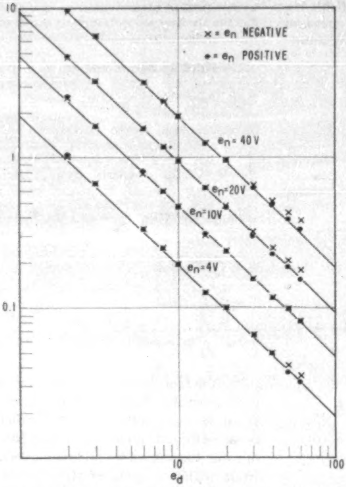


Fig. 6. Schematic of experimental circuit



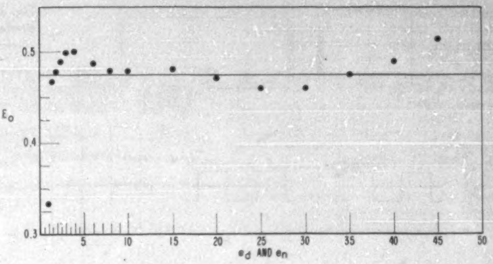


Fig. 7 (left). Average output voltage versus denominator voltage for various values of numerator voltage. Straight lines for function $E_0 = 0.475e_n/e_d$

Fig. 8 (above). Average output voltage for a constant ratio of input voltages

drop will be of importance unless it is two orders of magnitude less than e_d. Therefore, if i can be held small and constant, the effect of the iR drop will become unimportant for large values of e_d. This denominator circuit current i consists of two parts. One is the magnetizing current of the core and the other is the base current required by the transistors. A current limiter such as a biased rectifier in the base circuit of the transistors will keep this latter part of i small and constant. The magnetizing current component will be fairly constant for low driving voltages. However, as the core gets driven from saturation to saturation faster and faster, the magnetizing current requirements begin increasing somewhat exponentially. Eventually the iR drop again becomes of the same order of magnitude as e_d even though e_d is being increased. Therefore iR is of the same order of magnitude as e_d, both at low and high voltages.

If sinusoids are used as the input signals, the low limitation can be largely removed. This is accomplished by inserting a battery in the base circuit of the transistors so that the transistors will not conduct until the core voltage becomes considerably positive. Such a battery will affect the time in the cycle when the transistors fire, making the beginning of conduction much later in the cycle for small values of e_d. For large values of e_d, the battery affects this time of conduction only to a small extent. Part of the function of the 1.5-volt battery, Fig. 6, is to accomplish this. Varying the voltage divider will affect the value of the quotient at low values of e_d but leaves it nearly fixed for high values. This provides a method of adjusting the output to give a correct quotient at two points as a small denominator point can be adjusted to agree with a large denominator point.

Experimental Results

To determine the validity of the theory presented in the previous sections, a breadboard model of the divider was assembled. The circuit as built is shown in Fig. 6. The core is 1-mil *4168* Deltamax and the transistors are *2N44* PNP junction transistors. The 60-cycle commercial sinusoidal power was used as the inputs. Base current i_b was limited to 27 milliamperes by the biased rectifier in series with the core secondary. The 1.5-volt battery had the function of bucking out the minimum drop across the rectifier so that the transistors would not be conducting when there was no voltage on the core windings. In addition, this battery performed the corrective function described in the previous section.

Fig. 7 shows a log-log plot of the output as a function of e_d for various values of e_n. The straight lines are the values that would be obtained in a perfect divider if the output equaled 0.475 e_n/e_d. It can be seen that for e_d between 2 and 20 volts all values are within ±2½ per cent of full scale readings. The lower limit on the denominator voltage is near 2 volts as below this value the core will not be taken from saturation to saturation. At large values of the denominator voltage, increases in core magnetizing current begin introducing inaccuracies. Fifty volts is

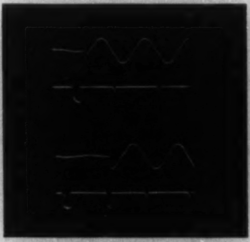


Fig. 9. Change in output voltage caused by a change in denominator voltage from 3 to 30 volts, numerator held constant

Top trace of each pair—Denominator voltage
Bottom trace of each pair—Output voltage

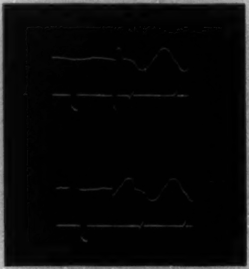


Fig. 10. Change in output voltage caused by a change of denominator voltage from −3 to +30 volts

Top trace of each pair—Denominator voltage
Bottom trace of each pair—Output voltage

the upper limit for the numerator voltage as above this the transistors break down.

These limitations are probably better shown in Fig. 8. Here the numerator and denominator voltages were held equal and together varied over a wide range. In a perfect divider the output would remain a constant. In Fig. 8 the variation is less than $2^{1}/_{2}$ per cent of full scale readings over the range from 1.5 to 40 volts.

Figs. 9 and 10 show the dynamic response of the divider. It should be noted that even when input wave forms undergo a transient complete response to a step, change of denominator is always achieved within $1^{1}/_{2}$ cycles of the change. In the bottom pair of traces of Fig. 9 it can be seen that there is, in fact, an instantaneous response to a change in magnitude of the denominator. Changes of the numerator, of course, show up equally fast.

Conclusion

It has been shown that with simple circuitry a dividing circuit with $\pm 2^{1}/_{2}$-per-cent accuracy can be built. The equipment is light and can be built compactly. Temperature should not be a major problem. According to the manufacturers' specifications, for instance, Deltamax varies in height by only 3 per cent for temperature variations from −60 to +180 degrees Fahrenheit. The variation in loop width is greater, but not of importance, if the impedance of the denominator voltage source is low. Temperature changes in the transistors are not as critical as in the usual applications as they are being used simply as switches.

Although the impedance of the denominator source must be low, this source is not required to deliver power. Before saturation the denominator circuit of Fig. 6 demands no more than 40 milliamperes. After saturation a current limiter such as two biased rectifiers can be used to limit the current if necessary.

Although the trial unit was operated at 60 cycles, there is no reason why such a device should not operate equally well up to frequencies of 1 kc. Finally the numerical values of accuracy presented should be thought of as typical accuracies rather than maximum accuracies as the circuit built is only a preliminary model. It is the hope that this unit will prove useful where a simple rugged device with reasonable accuracy is desired.

References

1. JUNCTION TRANSISTORS USED AS SWITCHES, R. L. Bright. *AIEE Transactions*, vol. 74, pt. I, March 1955, pp. 111–20.

2. TRANSISTOR DEMODULATOR FOR HIGH-PERFORMANCE MAGNETIC AMPLIFIERS IN A-C SERVO APPLICATIONS, R. O. Decker. *Ibid.*, pp. 121–23.

◆

Discussion

K. Chen and R. O. Decker (Westinghouse Electric Corporation, Pittsburgh, Pa.): The author is to be congratulated for making effective use of the new switching transistor techniques. The switching transistor should find wide applications in the field of analogue computation and control. Since the initial development of switching transistor circuitry, we have been working on analogue computing techniques utilizing the unique properties of switching transistors. To date, many schemes have been developed including an analogue divider which uses the same types of components but different circuitry from that which is described by the author.

One major difference between our circuitry and that of the author is the use of a variable-frequency magnetic-coupled oscillator. Controlled by the square-wave output of the oscillator, switching transistors are used which have either a driving or a blocking voltage on their bases. Thus, the leakage current through the switching transistors, when they are supposed to block, can be kept very small even at high temperatures. This is not the case with the circuit in Fig. 5 since the voltages across the emitter and base of the transistors are essentially zero when the core saturates. It is our experience that the leakage current through the switching transistors with zero blocking voltages can be quite appreciable, expecially at high temperatures.

D. H. Schaefer: It is in general necessary to have a bias voltage such as the 1.5-volt battery illustrated in Fig. 6. Such a bias will provide a blocking voltage on the transistor bases after saturation of the core.

The Application of Perpendicularly Superposed Magnetic Fields

ROBERT A. HEARTZ
ASSOCIATE MEMBER AIEE

HERBERT BUELTEMAN, JR.
NONMEMBER AIEE

Synopsis: The high permeabilities combined with high resistivities of recently developed ferrites makes practical a number of saturable reactor devices whose operation is based on the superposition of perpendicular fields. This paper presents a core configuration whereby the two perpendicular fields can be superposed and a number of applications for this type of reactor. Of particular interest is a perpendicular field low-level modulator.

THE characteristics of a ferromagnetic material which is magnetized in two perpendicular directions were discussed by Beck and Kelly[1] in 1948. This paper described how the core losses and permeability of the magnetic material in one direction is varied by superposing a field in a direction perpendicular to the first.

As an illustration of one means of obtaining this perpendicularity of fields in a ferromagnetic material, consider the core and coil configuration in Fig. 1. The core is a hollow torroid which incloses an annular winding and which is wound toroidally with a second coil. The magnetic fields produced by these windings are perpendicular and the mutual coupling between the windings is zero. Let it be assumed that a large direct current flows in the annular coil so as to saturate completely the core. Now, if a small field (a field which is not strong enough to reorient the magnetic domains) is applied by the toroidal winding, then the core as seen by this winding will appear as an air core. The core permeability will be unity and the core losses will be zero. If the large annular d-c field is removed, the core as seen by the toroidal winding will have its normally high permeability and losses. These phenom-

---

Paper 55-544, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 2, 1955; made available for printing May 9, 1955.

ROBERT A. HEARTZ and HERBERT BUELTEMAN, JR., are with the Minneapolis-Honeywell Regulator Company, Minneapolis, Minn.

The ferrite core material used in this investigation was Ferramic *H*, produced by the General Ceramics and Steatite Corporation.

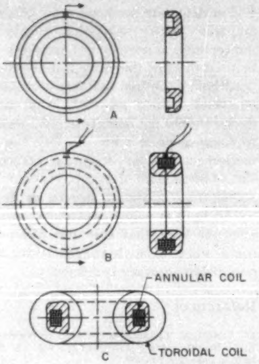


Fig. 1. Fabrication of cross-field modulator

ena are illustrated by the three hysteresis loops of Fig. 2. The largest figure is the normal 60-cycle hysteresis loop seen from the toroidal coil alone, whereas the smaller ones are those seen at two large values of annular current. The smallest hysteresis loop continues to show an appreciable area because of the relatively large alternating current used to obtain the loop.

Although these phenomena have been known for many years, only recently has a practical core material and core configuration been developed which allow the utilization of these characteristics. Solid magnetic metal cores have such a low resistivity that eddy-current losses are prohibitive, and there appear to be no practical laminated cores which are effective in two perpendicular directions. Powdered iron cores have such a low permeability that it is impractical to saturate them in any direction. A core material is needed which has high permeability as well as a high resistivity. The recently developed ferrites characterized by permeabilities on the order of 4,000 and resistivities up to 10^6 ohm-centimeters satisfy these conditions. Thus, the development of ferrite core materials has made practical a number of magnetic devices whose operation is based on the superposition of perpendicular magnetic fields. Possibly the most important of these devices is a low-level d-c magnetic modulator which uses the changing permeability caused by an a-c field to chop a perpendicular d-c signal field.

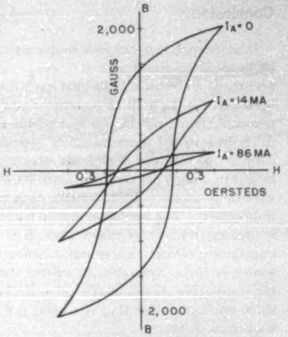


Fig. 2. Toroidal hysteresis loops as function of d-c annular current

The D-C Modulator

Magnetic modulation has often been considered for the conversion of d-c signals in the microvolt and millivolt regions. Magnetic modulators suggested heretofore have relied on the d-c signal biasing a magnetic core material so that the applied a-c excitation traverses an unsymmetrical hysteresis loop, giving rise to an induced voltage which contains a second harmonic component of the driving frequency. The induced fundamental is removed from the circuit by summing the outputs from two similar reactors. This type of modulator, frequently called the second-harmonic inverter, has been discussed in detail in past literature.[2,3] A modulator whose output is of the fundamental frequency can be obtained by supplying the two reactors with an a-c plus d-c field.[4] The unfortunate factor that limits the threshold in these units is that the two constituent reactors must be perfectly matched in differential permeability, maximum flux density, coercive force, and the temperature and pressure coefficients of these quantities, or an a-c output of the fundamental frequency can occur when the d-c signal is zero.

In the case of the fundamental output modulator, core mismatch will cause intolerable zero offsets and drifts. In the second-harmonic inverter, core mismatch will not produce zero offset or drifts, but will cause an excessively high noise level at zero signal. This places stringent requirements on the amplifier following the modulator. The ideal modulator is one in which the a-c output is derived directly from the d-c signal as in the case of the electromechanical chopper which converts the d-c signal to an a-c signal by varying the resistance of the d-c signal circuit. In this type of operation, the a-c output is inherently zero when the d-c signals is zero. The magnetic analogy to the chopper is a device which varies the reluctance of the d-c signal magnetic circuit. This can be accomplished by varying mechanically an air gap in the d-c signal magnetic circuit or by superposing an a-c field which is perpendicular to the d-c signal field and which saturates the material. The latter type of operation offers obvious advantages where life expectancy and insensitivity to external vibrations are of prime importance.

Consider the operation of the cross field

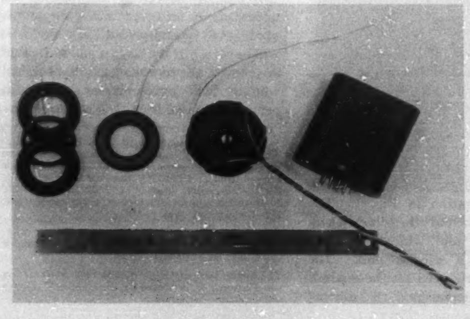


Fig. 3 (left). Assembly of a cross-field modulator

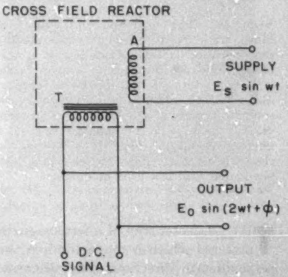


Fig. 4 (right). Second-harmonic cross-field modulator

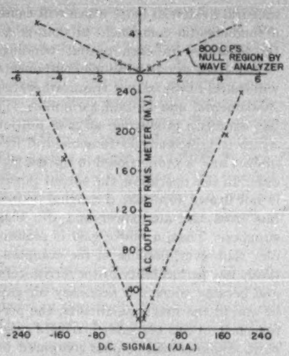


Fig. 5. Second-harmonic modulator transfer characteristics, 400-cycle operation

reactor in Fig. 1 when the core material is saturated by an alternating current flowing in the annular winding. Since there is no mutual inductance, there is no induced voltage in the toroidal winding. If the toroidal winding is now connected to a d-c source, a large flux will be set up normal to the toroidal coil when the annular field is passing through zero. However, this flux will be reduced to whatever value it would have for an air core at each saturation of the core by the annular field. Thus, the unit can be considered an electromagnetic chopper which develops an alternating voltage in the d-c signal winding whose magnitude is proportional to the impressed d-c signal and whose frequency is twice that of the saturating field. Fig. 3 shows the procedure of assembly of a cross-field reactor designed specifically for d-c modulation.

There are other core designs whereby two perpendicular fields can be superposed, but for optimum d-c modulation, the following conditions must be met:

1. To minimize zero drift due to hysteresis effects and to obtain a high modulation efficiency, the a-c supply must saturate the entire core volume.
2. To obtain good response times, induced currents within the core must be minimized.
3. To obtain high sensitivities, the core material must have a high permeability.
4. To minimize coupling between the windings, the windings must be accurately perpendicular.

The core of Fig. 1 is designed so the flux created by the annular winding sees a constant area in order that the entire core volume may be saturated. A ferrite core material was chosen to satisfy conditions 2 and 3. Since the air gap in the flux path of the annular coil appreciably decreases the effective permeability in this direction, the annular winding is used for the a-c excitation while the toroidal winding is used for the d-c signal.

A typical cross-field modulator circuit is shown in Fig. 4. Here the annular coil *A* is excited by a full-wave alternating voltage which saturates the core twice each cycle. Thus, the output signal will be a second harmonic of the supply frequency. The toroidal coil *T* is excited by the d-c signal. The threshold of this modulator is 10^{-14} watts (0.01 microampere) when operated into a second-harmonic selective system. Drifts due to changes in supply voltage and frequency (±10 per cent (%) of normal) and aging are less than the threshold signal. Some hysteresis effects were present when large signal excursions were applied, but below a ±10-microampere signal, these effects were entirely negligible. This modulator has a linear range (defined as the operating range where the ratio of d-c signal current to a-c output voltage is within ±10% of a constant) of 18,000. Modulation efficiency approaches 100%. The time constant is approximately 1.5 milliseconds. These values are typical of 12 experimental units tested, except that in two of the units a permanent zero offset of 0.04 microampere was observed. The output versus input characteristics of a typical unit are shown in Fig. 5.

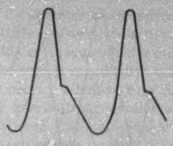


Fig. 7. Fundamental single-core modulator output wave form

In many applications a modulator whose output voltage is a second harmonic of the supply frequency is not desirable as it entails extra equipment in the overall system. The output frequency of the cross-field modulator can be fundamental if the supply field saturates the core only once each cycle. Thus, d-c biased excitation or a half-wave rectified current, as shown in Fig. 6, may be used to obtain the desired single saturation per cycle. The threshold signal of this modulator is 10^{-12} watts. When operating the modulator in this manner, an increase in threshold sig-

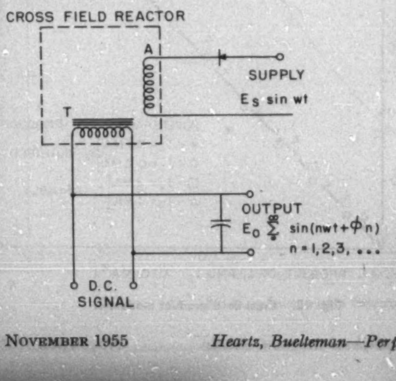


Fig. 6 (left). Fundamental cross-field modulator

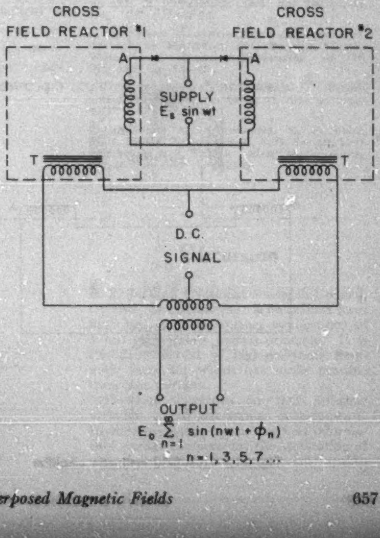


Fig. 8 (right). Two-core fundamental cross-field modulator

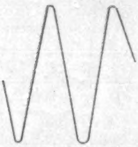


Fig. 9. Dual-core modulator output wave form at C = 0.2 microfarad

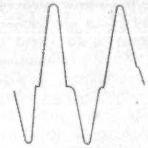


Fig. 10. Dual-core modulator output wave form at C=0.02 microfarad

nal may be expected because the noise pickup (due to capacitive coupling and inaccuracies in positioning of the two coils) is of the same frequency as the output signal. The wave form of the output signal for this type of operation is shown in Fig. 7.

The wave form of the fundamental output modulator can be improved by interconnecting two modulators as shown in Fig. 8 so that the modulators saturate on alternate half-cycles. Under these conditions, the even harmonics in the output wave form will cancel; Figs. 9 and 10. A further advantage of this circuit is that the output voltage is independent of the d-c signal source impedance. Because of this independence, the circuit can provide a substantial voltage gain when the output is operated into a high impedance such as a vacuum-tube amplifier. At 400 cycles-per-second operation, voltage gains of 1,000 have been realized.

The performance figures given in the preceding represent those of a typical unit under conditions practical for aircraft control systems, i.e., the supply frequency is 400 cycles per second, the supply variations are ±10%, harmonic content of the supply is greater than 1%, and the allowable ambient temperature range is −65 to 165 degrees Fahrenheit. Under rigidly controlled ambient and operating conditions, thresholds on the order of 10^{-16} watts should be attainable.

Other Applications

In modulator circuits where zero stability is of prime importance, the a-c output is derived directly from the d-c signal. This type of operation leads to better zero stability, but obviously the power gain cannot exceed one. A circuit whereby a power gain can be realized is shown in Fig. 11. The operation of this circuit is as follows: When the d-c signal is zero, the d-c bias is modulated by the a-c supply field and produces a second-harmonic output in each core. These second harmonic currents are subtracted in the output transformer. When a d-c signal current flows, it will cause the a-c supply field in each core to traverse an unsymmetrical hysteresis loop, which will cause a fundamental component of voltage to be generated in each output winding. Whereas the second harmonic components will cancel in the output transformer, the fundamental and all odd harmonics will be summed, producing an a-c output which is principally of fundamental frequency and is proportional to the d-c signal. In this operation, the output power is not drawn from the d-c signal source but from the a-c power and d-c bias supplies. Thus, a power gain is realizable. However, because of the comparatively low permeability of the ferrite core and because there is a necessary air gap in one of the magnetic circuits, the performance of this perpendicular-field magnetic amplifier will be poor compared to that of conventional magnetic amplifiers when operated at power frequencies. However, at higher frequencies, where ferrite cores will also be required in conventional magnetic amplifiers, the perpendicularfield amplifier may offer several advantages. It should be further noted that although the output power is basically drawn from the d-c bias source, an a-c power source is required to modulate the d-c bias field. Thus the unit inherently must consume more a-c power than it puts out.

While the application of the circuit shown in Fig. 11 as a magnetic amplifier at power frequencies is not very practical, the circuit does provide a unique and compact computer element. The a-c output of this circuit is directly proportional to the product of the d-c bias and the d-c signal. The graph of Fig. 12 illustrates

Vol. 7
Pt. 1
1955

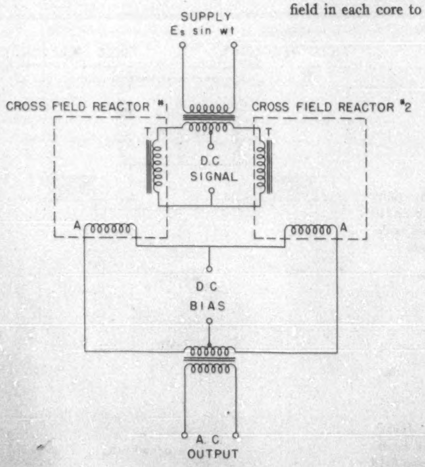


Fig. 11. Cross-field magnetic amplifier

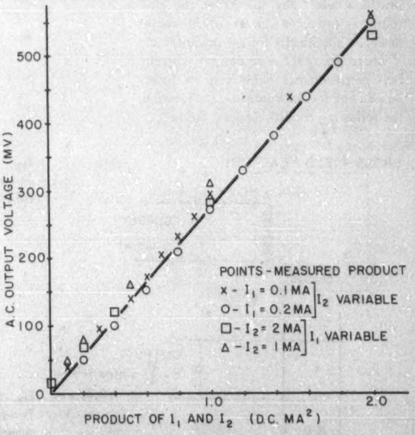


Fig. 12. Cross-field product modulator

this type of operation where the output voltage is shown as a function of the product of the d-c bias and d-c signal for a number of conditions. If the d-c bias is held fixed and the d-c signal is varied, then the ideal relationship between the output voltage and the product of the two direct currents should be linear. The dashed straight line in Fig. 11 illustrates this ideal condition. The reversal of either the d-c signal or bias will reverse the output phase, thus giving a polarity sensitive unit.

The fact that the permeability of one flux path is an inverse function of the normal field strength can be utilized to construct an electrically variable inductance. The inductance of the toroidal winding of the cross-field reactor is plotted as a function of annular current in Fig. 13. Such a unit could be used in sweep frequency oscillators, automatic-frequency control systems, variable filters, etc.

The cross-field reactor is similar to the cross valve as described by McCreary[5] except that the former has no core area which is not common to both fields. It is found that the cross-field reactor will perform the functions of the cross valve, i.e., when full-wave excitation is used, the toroidal winding can be tuned to give an output of the second or third harmonic of the excitation. Under half-wave excitation, proper tuning results in output voltages of one-half and one-third the frequency of the supply voltage. When the cross-field reactor is used for d-c modulation, it is often desirable to load the a-c output with a capacitor in order to improve the wave form or to shift the phase position. In this case, caution must be exercised to prevent the cross-valve type of operation. Generally, the d-c signal source impedance is sufficiently low to prevent cross-valve action.

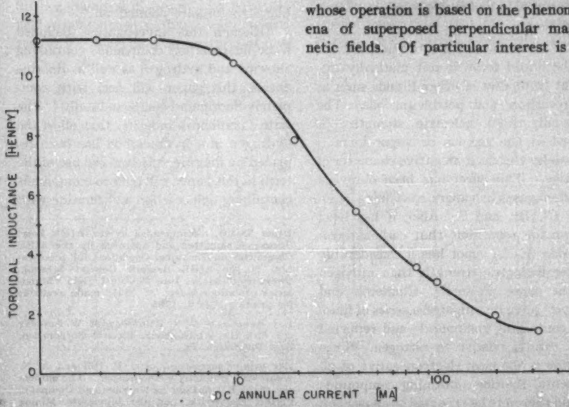


Fig. 13. Cross-field variable inductor

The hollow toroidal configuration is optimum for the foregoing application because of its closed flux paths in the core material. However, for some purposes, other constructions are more suitable. Consider the action of a cross-field reactor having the configuration used by Beck and Kelly[1] (a hollow pipe wound both toroidally and solenoidally) when saturated by an alternating current in the toroidal winding. An a-c output will occur in the solenoidal winding whenever a d-c axial field is present. Such a reactor might be used as a magnetometer, where the ambient fields of interest would supply the axial flux. This type of magnetometer would have the sensitivity of the conventional magnetor type but would be superior to it in that the output would have a minimum induced fundamental to be filtered out.

The magnetometer configuration might also be used to provide a current reference system by biasing the core with a small permanent magnet. The complete reference system would amplify and demodulate the a-c output and feed this signal back to the solenoidal winding in such a sense as to cancel the permanent magnet field.

Summary

The recently developed ferrites, which have high resistivities combined with high permeabilities, have made practical a number of saturable reactor devices whose operation is based on the phenomena of superposed perpendicular magnetic fields. Of particular interest is a low-level d-c inverter whose operation is based on the use of an a-c field to vary the permeability of a perpendicular d-c signal field. The two perpendicular fields are best obtained by a hollow toroidal reactor which has both an annular winding within the core and a toroidal winding. Using two cross-field reactors, a modulator can be made which has a large voltage amplification and an output which is independent of the signal source impedance. These features are very desirable in many applications where the signal source is of very low impedance, such as a thermocouple. These two reactors, with different external circuitry, can also be used as a unique product modulator where the output is proportional to the product of two d-c signals. A single cross-field reactor can be used as a means for obtaining an electrically variable inductor or as a frequency multiplier or divider. The latter operation is similar to that described by McCreary.[5] A cylindrical-core cross-field reactor offers some very attractive features when used as a magnetometer or a permanent-magnet biased reference source.

References

1. Magnetization in Perpendicularly Superposed Direct and Alternating Fields, F. J. Beck, J. M. Kelly. *Journal of Applied Physics*, New York, N. Y., vol. 19, June 1948, pp. 551–62.

2. The Fundamental Limitations of the Second Harmonic Type of Magnetic Modulator as Applied to the Amplification of Small D-C Signals, F. C. Williams, S. W. Nobel. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 97, pt. II, Aug. 1950, pp. 445–59.

3. The Study of a Magnetic Inverter for Amplification of Low Input Power D-C Signals, E. H. Frost-Smith. *Ibid.*, vol. 100, pt. II, Aug. 1953, pp. 362–75.

4. A Simple Magnetic Modulator for Conversion of Millivolt D-C Signals, Gunnar Wennerberg. *Electrical Engineering*, vol. 70, Feb. 1951, pp. 144–47.

5. The Magnetic Cross Valve, H. J. McCreary. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1868–75.

6. Airborne Magnetometers, E. P. Felch, W. J. Means, T. Slonczewski, L. G. Parratt, L. H. Rumbaugh, A. J. Tickner. *Electrical Engineering*, vol. 66, July 1947, pp. 680–85.

◆

Discussion

R. R. Jackson (National Research Council, Ottawa, Ont., Canada): The authors have described some interesting applications of a not very widely known principle. It is the development of high-resistance magnetic materials which has made possible these applications.

Previous literature on this principle includes some interesting developments. In the use of tape-wound cores as information storage components, the application of a perpendicularly superposed field makes possible the nondestructive sensing of the information.[1]

The authors' suggestions for a cross-field magnetometer have already been applied by T.M. Palmer in Australia.[2] Here a short piece of fine mumetal wire carries alternating current which provides an alternating circular field in the wire. If the wire is in an ambient steady field, voltage of twice excitation frequency is induced in a coil coaxial with the wire.

Work very similar to that of the authors is described briefly in reference 3, and an application of the principle to a clip-on ammeter for direct currents is shown in reference 4.

REFERENCES

1. NONDESTRUCTIVE SENSING OF MAGNETIC CORES, D. A. Buck, W. I. Frank. *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 822–30.

2. A SMALL SENSITIVE MAGNETOMETER, T. M. Palmer. *Proceedings*, Institute of Electrical Engineers, London, England, pt. II, Oct. 1953, p. 545.

3. A TRANSDUCER USING CROSSED MAGNETIC FIELDS, E. H. Frei, S. Shtrikman, D. Treves. *Bulletin*, Research Council of Israel, Jerusalem Israel, Mar. 1954, p. 443.

4. A CLIP-ON MILLIAMPEREMETER FOR D.C. CURRENTS, E. H. Frei, D. Treves. *Ibid.*, p. 444.

Electric Breakdown of Perfluorocarbon Vapors and Their Mixtures With Nitrogen

L. J. BERBERICH
FELLOW AIEE

C. N. WORKS
MEMBER AIEE

E. W. LINDSAY
ASSOCIATE MEMBER AIEE

Synopsis: This investigation was concerned with a study of the breakdown behavior of the higher molecular weight perfluorocarbon vapors and mixtures of these with nitrogen. Impulse and 60-cycle data were obtained under both uniform and nonuniform field conditions. The results are compared with those for sulfur hexafluoride, nitrogen, and oil. It was found that these perfluorocarbon vapors have surprisingly high dielectric strength not only in the pure form but also when mixed with nitrogen. Values of dielectric strength higher than for sulfur hexafluoride are easily obtainable at the same pressure. Values equal to those for oil can be obtained at relatively low pressures under certain conditions. The results are briefly discussed in terms of modern gas breakdown theory. The unusually high dielectric breakdown combined with complete nonflammability, inertness, and excellent thermal stability of the perfluorocarbons suggest that they should find applications in the insulation field.

THE gaseous and liquid perfluorocarbons, developed during World War II are among the more recent synthetic materials which show promise of finding applications in the electric insulation field. Solid fluorocarbons such as resinous polyfluorocarbons sold under the trade names of Teflon, Kel-F, and Fluorothene are well known and have found a considerable number of specialized uses not only in the electrical industry but also in many other industries. The lower molecular weight gaseous fluoro-chloro-carbons are also well known. These materials have been marketed under the trade name of Freon and have been used for many years as refrigerants, coolants, and in some cases as dielectrics. The higher molecular weight gaseous and liquid perfluorocarbons came into being through the pioneer work done by The Johns Hopkins University, Pennsylvania State University, and Purdue University, and also by E. I. duPont de Nemours and Company and more recently by the Minnesota Mining and Manufacturing Company. The perfluorocarbons are compounds of carbon and fluorine which show quite different properties from their hydrocarbon counterparts. Perhaps their most outstanding properties are their nonflammability, inertness, excellent thermal stability, and high dielectric strength, particularly in the gaseous or vapor form. These properties suggest that this type of fluorocarbon should find applications as electric insulation.

The dielectric strength of fluorocarbons in the liquid form[1] is not markedly different from that of other liquids such as hydrocarbon and petroleum oils. The unusually high dielectric strength observed in the gaseous or vapor form is caused by the electronegative character of fluorine. This effect has been observed in other gases or vapors containing halogens Cl, Br, and F. Also, it has been known for some time that carbon tetrachloride (CCl_4) vapor has a considerably higher dielectric strength than nitrogen at the same pressure.[2] Charleton and Cooper[3] have investigated a series of halogen containing compounds and reported their results relative to nitrogen. Since fluorine is the most electronegative of all elements, fluorine-containing compounds should therefore be expected to show even higher dielectric strength than most other electronegative gases. It has already been shown that this is true for sulfur hexafluoride[4,5] and other fluorine containing gases.[6] More recently, Wilson et al.,[7] have investigated the lower molecular weight straight-chain perfluorocarbons through perfluoropentane (C_5F_{12}) and found high dielectric breakdown values. Thus, it is apparent that fluorine-containing compounds are particularly effective in trapping electrons in the early stages of the gaseous breakdown process and thereby increasing the voltage at which final breakdown takes place.

The primary purpose of this paper is to report the original work on the dielectric breakdown of higher molecular weight peruflorocarbon vapors. This work was started in 1947 at the Westinghouse Research Laboratories. Straight-chain and cyclic perfluorocarbon fluids with boiling points up to 200 degrees centigrade (C) were investigated. Vapors of some of these perfluorocarbons alone and when mixed with nitrogen showed surprisingly high dielectric breakdown values. The breakdown mechanism in electronegative gases and vapors is discussed briefly. Possible relationships between breakdown voltage and physical properties or molecular structure are pointed out.

Although this investigation included some fluorocarbon compounds containing chlorine and hydrogen as well as fluorine atoms, this paper will deal with completely fluorinated compounds only. The term "perfluoro" indicates that all of the hydrogen in a hydrocarbon has been replaced by fluorine. Hence, the use of the term in this paper will refer to compounds containing only carbon and fluorine with

---

Paper 55-493, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 25, 1955; made available for printing May 3, 1955.

L. J. BERBERICH, C. N. WORKS, and E. W. LINDSAY are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

The helpful suggestions of C. F. Hill and R. N. Wenzel are gratefully acknowledged. The authors also thank E. I. duPont de Nemours and Company, Purdue University, and the Minnesota Mining and Manufacturing Company for supplying samples of the materials investigated.

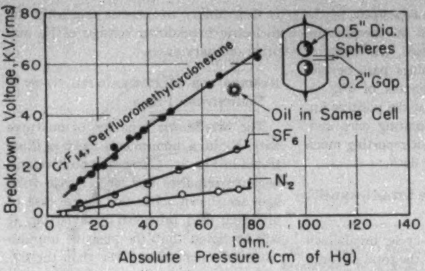


Fig. 1. 60-cycle breakdown voltage versus absolute pressure for N_2, SF_6, and C_7F_{14} using spherical electrodes in 4-inch-diameter glass cell

the exception that a few may also contain oxygen and nitrogen atoms. The presence of nitrogen or oxygen in the molecule, however, has little or no effect on the physical or electric properties of the compounds.

Experimental Results

The perfluorocarbon compounds investigated are listed in Table I which also gives some of their physical properties such as molecular weight, density, boiling point, and vapor pressure. It will be observed that the boiling points of these compounds range from −128 to over 200 C and molecular weights from 88 to over 600. The 60-cycle dielectric breakdown relative to N_2 of each compound is given in the last column for comparative purposes. The compounds in Table I are arranged in ascending order of breakdown strength. The detailed experimental results on a number of pure perfluorocarbon compounds in the form of curves are given first. These data are further classified in accordance with the field configuration employed in making the tests. Then data on mixtures of perfluorocarbons with nitrogen and with sulfur hexafluoride are treated in the same manner. These results are compared to those for the pure compounds. Impulse and 60-cycle voltages are applied although both were not always used in the same series of tests. All impulse tests were made with the standard 1½x40-microsecond wave form. Since many of the compounds investigated are liquids at room temperature, the desired vapor pressure was obtained by controlling the temperature of the test apparatus.

Results for Pure Compounds in Uniform Field

Nearly uniform field conditions were obtained by using a 0.2-inch gap between spherical electrodes enclosed in a 4-inch-diameter glass cell. Fig. 1 gives 60-cycle breakdown data as a function of pressure of perfluoromethylcyclohexane (C_7F_{14}) compared with that for other gases. The behavior of C_7F_{14} is typical of that for other compounds in this class. These data show that the breakdown voltage of this compound has a considerably higher value than that for sulfur hexafluoride or for nitrogen and that at one atmosphere pressure it has about the same breakdown value as oil. As will be shown later, this comparison with oil, however, is not valid under nonuniform field conditions.

Many of the 60-cycle tests were made with a concentric glass tube ozonizer apparatus. This test apparatus proved to be particularly convenient because the glass tubes served not only to confine the vapors but also to limit the discharge current, thus avoiding electrode damage. This electrode arrangement is shown in Fig. 2. It should be pointed out that there are two layers of glass between the electrodes. The results are reported in terms of the voltage at which the discharge appeared with the various gases or vapors in the space between the glass tubes. In general, concentric electrodes produce a nonuniform field. Since, however, the radius of the inner electrode is not greatly different from that of the outer electrode, the field was only very slightly divergent and, hence, approximates the uniform field case. The data of Fig. 2 obtained in the ozonizer compare the 60-cycle results on a number of perfluorocarbon vapors with sulfur hexafluoride and nitrogen. It will be observed that, with the exception of perfluoromethane (CF_4), the perfluorocarbons have larger values of breakdown voltage than sulfur hexafluoride, and all have higher values than nitrogen. Again it is apparent, that C_7F_{14} has a much higher value of breakdown voltage than sulfur hexafluoride. A study of Fig. 2 shows that, starting with the small simple molecules such as CF_4 which is a gas at one atmosphere even at temperatures below −100 C, then proceeding to the larger molecules which are liquids having boiling points of over 80 C, the dielectric breakdown increases markedly. It may also

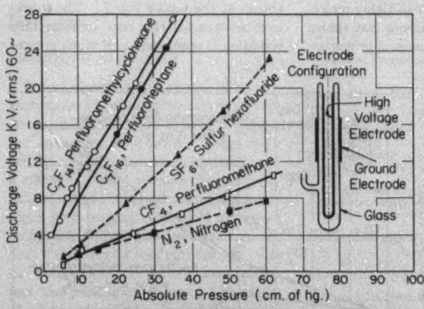


Fig. 2. Discharge voltage of a number of fluorocarbon gases and vapors compared to that of several other gases and vapors

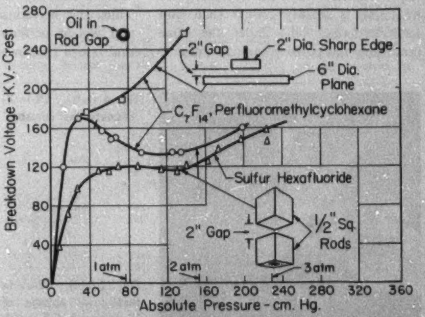


Fig. 3. Positive impulse breakdown voltage of C_7F_{14}, SF_6, and transformer oil versus absolute pressure

be observed from Fig. 2 that the breakdown voltage of the straight-chain perfluorocarbons, e.g., perfluoroheptane, is not markedly different from that of the cyclic compound C_7F_{14}.

Results for Pure Compounds in Nonuniform Field

The curves of Fig. 3 show the positive impulse breakdown voltage of C_7F_{14} as a function of pressure compared with sulfur hexafluoride. These data also show the effect of the two electrode configurations. The rod gap was mounted in a 12-inch-diameter steel tank shown in Fig. 4. These sharp-edged electrodes produce a very highly divergent field. For this condition, at pressures between 1 and 3 atmospheres, it is evident from Fig. 3 that the positive impulse breakdown of C_7F_{14} is only slightly greater than that of SF_6 and considerably less than that of oil.

The other electrode configuration, a sharp-edged disk to plane shown in Fig. 3, represents a somewhat less severely divergent field than for the rod gap. Even this gap is generally more severe than encountered in practice since the high-voltage parts usually have at least a slight radius on the edges. Note that this sharp-edged disk-to-plane geometry results in a very marked improvement over rod-gap breakdown. At 2 atmospheres' pressure, the positive impulse breakdown is equal to that of oil. Thus, it is apparent from the data in Fig. 3 that the breakdown voltage of the perfluorocarbon vapors is extremely sensitive to electrode configuration.

The negative impulse breakdowns for these compounds were found to be considerably higher than the positive impulse values. For example, with the rod gap of Fig. 3, at a pressure of only 50 centimeters of Hg absolute, the negative impulse breakdown of C_7F_{14} is about 260 kv. This is slightly greater than that for transformer oil. Under the same conditions the positive impulse value is only 160 kv or only 61.5 per cent of the negative impulse value. As the pressure of the vapor is increased, the negative impulse breakdown values pass beyond the range of the ordinates of Fig. 3. Thus, it is apparent that the positive impulse values are the limiting ones and, hence, the reason for not reporting much of the negative impulse data.

Results for Mixtures with N_2 and SF_6 in a Uniform Field

Fig. 5 gives the 60-cycle breakdown voltage as a function of the total pressure for mixtures of C_7F_{14} with N_2 compared with mixtures of the same compound with SF_6. This comparison is made for uniform field conditions. Liquid perfluorocarbon was introduced into the cell, and increase in pressure of the vapor was obtained by raising the temperature of the cell. Several temperatures are indicated on the curves. When the first small amount of a perfluorocarbon vapor is added to N_2 a rapid increase in breakdown voltage takes place, as indicated by the dashed portion of the lower curve. The breakdown voltage then increases linearly with increasing pressure. In contrast to this, no such rapid rise in breakdown values are observed for the first small additions of perfluorocarbons in SF_6. It is somewhat surprising to note that the breakdown values for mixtures of perfluorocarbons with SF_6 are only slightly higher than for the corresponding mixtures with N_2. It should also be observed from Fig. 5 that under uniform field conditions a breakdown voltage equal to that for oil is obtained for a pressure of 118 centimeters of Hg absolute in the case of C_7F_{14}-N_2 mixture. It should also be pointed out that for uniform field conditions the positive impulse, the negative impulse, and the 60-cycle crest values have practically the same magnitude. Thus, this investigation of higher molecular weight perfluorocarbons has shown that one of their most unusual properties is their ability to increase markedly the dielectric breakdown voltage of N_2 and other ordinary gases.

Results for Mixtures with N_2 in a Nonuniform Field

The breakdown behavior of mixtures with N_2 in a nonuniform field is illustrated in Fig. 6. Curves for positive and negative impulses and for 60-cycle voltages are shown. For this severe case of divergent field, produced by a rod gap, it will be noted that the positive impulse values are somewhat lower than the 60-cycle crest values, whereas the reverse is true for the negative impulse values. Fig. 6 reveals that for mixtures the increase in dielectric breakdown voltage over that for N_2 alone is quite striking with relatively small additions of C_7F_{14}. For example, 0.12 pound per cubic foot of C_7F_{14} added to N_2 increases the 60-cycle dielectric breakdown by a factor of about 3.5. This amount of vapor produces saturation of N_2 at 25 C. Under these conditions, the partial pressure of the perfluorocarbon is only about 12 centimeters of Hg. It will also be observed that the gain is least for positive impulse but, even in this case, a concentration of only 0.15 pound of vapor per cubic foot of gas space results in an increase in breakdown voltage by a factor of 3. Under the same conditions, the negative impulse breakdown is increased by a factor of 4. This concentration of vapor occurs at a temperature slightly above 25 C.

Fig. 7 shows the effects of two electrode configurations on the positive impulse breakdown of mixtures of C_7F_{14} with N_2. It also gives a comparison between the results for mixtures and the pure compound for the rod-gap electrode shape. The breakdown between plates having rounded edges at a total pressure of 1.5 atmospheres is shown by a single solid square at the top of Fig. 7. The sparkover voltage for this near uniform field condition is observed to be much greater

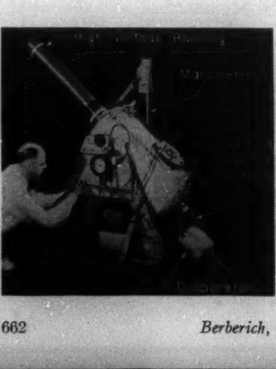


Fig. 4 (left). 12-inch-diameter test chamber

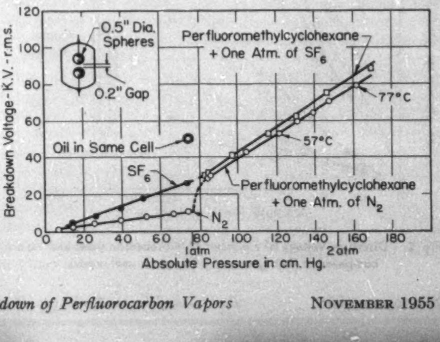


Fig. 5 (right). 60-cycle breakdown voltage of mixtures of C_7F_{14} with SF_6 and N_2 as a function of pressure

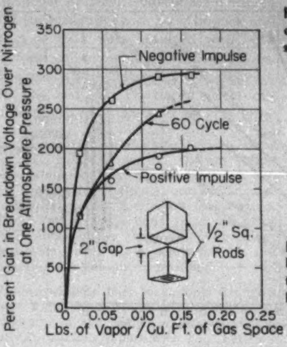
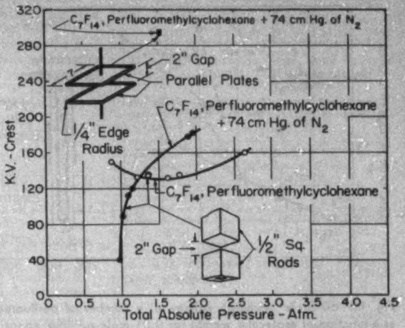


Fig. 6 (left). Increase in breakdown voltage of N_2 by adding small amounts of C_7F_{14} vapor

Fig. 7 (right). Positive impulse breakdown of a 2-inch gap for two electrode configurations as a function of pressure for C_7F_{14} and a mixture of C_7F_{14} with N_2



than for the same mixture tested under the severely divergent test conditions of the rod gap. It is therefore apparent that considerable benefit is derived by rounding the edges of the high-voltage parts, thus reducing the degree of nonuniformity of the field.

Comparison of the two curves of Fig. 7 also shows that for the rod gap only small additions of perfluorocarbon vapor to N_2 are required to equal and exceed the impulse breakdown values of the pure vapor. Other comparisons can also be made for 60 cycles by relating the data of Figs. 1 and 5. From these it can be determined, for the uniform field case also, that a given breakdown voltage can be obtained with roughly two-thirds of the amount of C_7F_{14} when mixed with N_2 that is required when the pure vapor is used alone.

This behavior of N_2-perfluorocarbon mixtures, especially in the nonuniform field case, was one of the most surprising results of this investigation. This effect was discovered by Berberich and Works[8] a number of years ago.

Discussion

BREAKDOWN MECHANISM

Before discussing the results of this investigation in terms of breakdown mechanism, it should prove helpful to review the gaseous breakdown theory very briefly. Electronegative elements exert a strong influence on such a breakdown. At the outset, it appears desirable to explain why several test conditions were employed. The data definitely indicate that the electrode configuration, the rate of voltage rise, and polarity strongly affect the breakdown values. The effects of all these variables must be known in any practical or theoretical study. The ultimate goal of all theoretical studies of dielectric behavior is to obtain constants for a given material which will permit computation of its breakdown properties under all conditions of use. Some recent contributions to the general gaseous breakdown theory have been made by Kihara[9] and Abdelnabi.[10] The breakdown of an electronegative gas in a uniform field has been treated by Geballe and Reeves.[11] Such investigations show that physicists are making progress toward the goal of the theoretical calculation of breakdown potential in all gases. However, the present state of knowledge limits the discussion of the breakdown processes in gases to a qualitative basis.[12]

The initiation of breakdown processes for ordinary gases in a d-c field for an extremely divergent field such as a point-to-plane electrode system and with the point positive may be explained as follows: Electrons are attracted to the point from the weaker parts of the field, leaving a positive-ion space charge which grows outward into a streamer with a concentrated electric field at its tip. These streamers form few branches because most of the electrons join the advancing streamer at its tip. Now, when the point is negative, streamers also form, but electrons from the surrounding region are repelled. The electrons following up the streamer from the negative electrode point spread out into numerous weak branches, scattered in all directions toward the weaker part of the field. It is this process of concentrated streamer formation that results in the observed lower breakdown from a positive point. It also explains why the positive impulse strength for the rod gap reported in this paper is considerably lower than the negative impulse strength.

The breakdown mechanisms for electronegative gases in a nonuniform field have been treated in an earlier paper.[13] The increase in the breakdown strength of gases containing electronegative mole-

Table I. Physical Properties and Breakdown Voltage of Compounds Investigated

Name of Compound	Molecular Formula	Molecular Weight	Liquid Density at 20 C	Refractive Index at 20 C	Boiling Point, C	Vapor Pressure at 25 C Millimeters, Hg	Relative 60-Cycle Breakdown at 30 Centimeters, Hg
Nitrogen	N_2	28	0.808 at −196 C		−196	gaseous	1
Perfluoromethane	CF_4	88	1.96 at −184 C, gas 0.60		−128	gaseous	1.25
Perfluoropropane	C_3F_8	188	1.45, at −38 C, gas 0.61		− 38	gaseous	2
Sulfur hexafluoride	SF_6	146	1.91		− 63.8	gaseous	2.5
Perfluoroheptane	C_7F_{16}	388	1.733	1.2618	82	110	5.5
Perfluorotriethylamine	NC_6F_{15}	371	1.749	1.262 at 25 C	71	130	6
Perfluoromethylcyclohexane	C_7F_{14}	350	1.7994	1.2815	76	110	6
Perfluorotoluene	C_7F_8	236	1.660	1.3664	102	23	7
Perfluorodibutyl ether	$C_8F_{18}O$	454	1.717 at 25 C	1.26	100	33	7.5
Perfluorotri-n-butylamine	$NC_{12}F_{27}$	671	1.88	1.29 at 25 C	178	2.7	8
Perfluorodimethylcyclohexane	C_8F_{16}	400	1.8503	1.2896	101	40	8.5
Perfluorophenanthrane	$C_{14}F_{24}$	624	2.019	1.3315	205		10

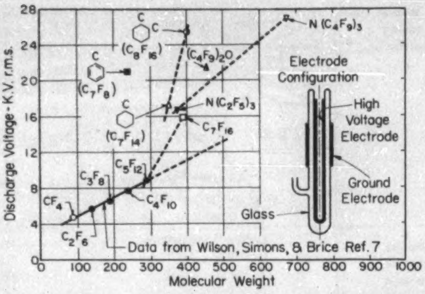


Fig. 8. 60-cycle discharge voltage of a number of perfluorocarbons as a function of molecular weight for 0.01-mole-per-liter concentration

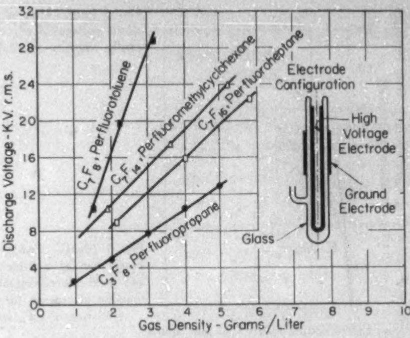


Fig. 9. 60-cycle discharge voltage of a number of perfluorocarbons as a function of density

cules may be explained very briefly as follows: Because of the great affinity of electronegative atoms for electrons, many free electrons are captured before they can reach ionizing velocities. Thus streamer formation is inhibited. The slow negative ions formed through electron capture also reduce the positive space charge in the streamer, thus making necessary a higher voltage before the streamer can cross the gap. In common with other gases, electronegative gases may also extract energy from electrons without actual capture through inelastic collisions. Inelastic collisions are more probable when large molecules are present. Such collisions are important in this discussion because many of the perfluorocarbon vapors investigated are large molecules.

BREAKDOWN OF PERFLUOROCARBONS IN RELATION TO PROPERTIES AND STRUCTURE

As was pointed out previously, Table I lists the compounds in ascending order of breakdown relative to N_2. This table shows that the breakdown strength at the same pressure increases in general with increasing molecular weight or molecular size. However, it also shows that there are exceptions. Perhaps the outstanding exception is perfluorotoluene (C_7F_8). This is the only cyclic aromatic compound investigated. It possesses a conjugated system of double bonds in the 6-member ring. This resonating system of double bonds has a high electronic polarizability, as indicated by the high value of refractive index also given in Table I. This could very well explain why it is more effective in absorbing energy from electrons.

Fig. 8 shows the results for some of the compounds plotted against molecular weight. To make the results strictly comparable, the same number of molecules must always be between the electrodes and, hence, the results of Fig. 8 were obtained for a vapor concentration of 0.01 mole per liter in every case. Starting with the lower molecular weight straight-chain compounds investigated by Wilson, Simons and Brice,[7] it will be observed that a straight line is obtained through the C_5F_{12} compound. This relation, however, does not hold for the higher molecular weight straight-chain compounds since the value for perfluoroheptane (C_7F_{16}), indicated by an open square in Fig. 8, lies considerably above the dashed extention of the solid curve. It should be pointed out that the data on the lower molecular weight compounds obtained in the ozonizer apparatus agree reasonably well with those obtained by Wilson and co-workers.

The perfluoro-ether $(C_4F_9)_2O$ and the perfluoroamines, NC_6F_{15} and $NC_{12}F_{27}$, as well as (C_7F_{16}) all behave similarly, since the results for all of them fall close to the dashed straight line at the upper right of Fig. 8. Apparently the presence of the oxygen atom in the ether compounds and the N_2 in the amine compounds do not alter the breakdown behavior.

The cyclic compounds whose skeletal and molecular formulas are given in Fig. 8 all show exceptionally high values of breakdown. This is particularly true of C_7F_8 which was discussed previously. Reasons for the unusual behavior of these cyclic compounds, other than those already mentioned, are not known.

Wilson, et. al.,[7] have found that the breakdown strength of the straight-chain compounds through C_5F_{12} is approximately linearly related to the density of the vapor. To determine whether such a relationship also holds for the higher molecular weight compounds, the breakdown data were plotted against density in Fig. 9 for a few representative compounds. This plot indicates that there is a definite linear relationship with density for any individual compound, but there was no linear relationship with density among the different compounds. If there were a linear relationship with density among the compounds, all of the compounds would fall on a single straight line. It is quite obvious from Fig. 9 that this is not true. Again the cyclic compounds and particularly C_7F_8 stand out as having unusually high values for a given density.

It is thus apparent that no single property of these compounds was found to be directly related to dielectric breakdown. A previous attempt was made to relate breakdown to molecular weight in hydrocarbon vapors and only a loose relationship was found.[14] It can be said, however, in general, that the larger the molecule the higher the dielectric strength. Many factors influence the breakdown process. It appears that the details of molecular structure are very important in determining breakdown strength. The behavior of C_7F_8 with its system of resonating conjugated double bonds is an outstanding example of the influence of structure. In spite of the fact that C_7F_8 contains considerably less fluorine than its saturated counterpart C_7F_{14}, its breakdown strength is higher. The resonating system of double bonds and the associated higher electronic polarizability apparently facilitate the absorption of the energy from free electrons in the early stages of the breakdown process.

Summary and Conclusions

The investigation involves the study of the breakdown behavior of the higher molecular weight perfluorocarbon vapors and their mixtures with N_2. Impulse and 60-cycle data were obtained with both uniform and nonuniform field conditions. The results were compared with those for SF_6, N_2, and oil. The major conclusions to be drawn from this work are as follows:

1. Comparison of the results with those for SF_6 indicates that breakdown values approximately four times those for SF_6 are possible for the pure form of the high molecular weight compound investigated. Breakdown strengths equal to oil and six or more times higher than N_2 appear to be realizable under some conditions by use of C_7F_{14} at atmospheric pressure.

2. Surprisingly high breakdown strengths were found for mixtures of relatively small amounts of the higher molecular weight perfluorocarbon vapors with N_2. For 0.15 pound per cubic foot of C_7F_{14} added to N_2 at 1-atmosphere pressure, the positive impulse breakdown is increased over that of N_2 alone by a factor of 3.

3. Over a considerable range of pressure, the breakdown values of mixtures with N_2 were greater than for the pure vapor at the same pressure. The practical implication of this is that less perfluorocarbon is required to obtain a given dielectric strength by use of mixtures with N_2 than by use of the pure vapor.

4. The breakdown values for both the pure perfluorocarbons and their mixtures with N_2 were found to be strongly affected by the field configuration of the test electrodes which varied from spheres to sharp-edged rods. Considerably higher values, however, were obtained with a sharp-edged rod to plane than for a pair of identical sharp rods.

5. Significantly higher values of breakdown were found for negative impulse waves than for positive impulse waves under nonuniform field conditions for both the pure compounds and the mixtures.

6. It appears that the breakdown strength of the perfluorocarbons is influenced not only by their molecular weight but also by their chemical structure.

7. The perfluorocarbons, in addition to having high-dielectric strength, have other desirable properties for electrical application such as flameproofness, excellent thermal stability, chemical inertness.

References

1. Some Fluorinated Liquid Dielectrics, **N. M. Bashara.** *AIEE Transactions*, vol. 72, pt. I, March 1953, pp. 79–85.

2. Effect of CCl_4 Vapor on the Dielectric Strength of Air, **M. T. Rodine, R. G. Herb.** *Physical Review*, New York, N. Y., vol. 51, 1937, pp. 508–11.

3. The Dielectric Strength of Insulating Fluids. Part 1: Electric Breakdown Strengths of Certain Gases and Gas-Vapor Mixtures, **E. E. Charleton, F. S. Cooper.** *General Electric Review*, Schenectady, N. Y., vol. 40, 1937, pp. 438–42.

4. The Effect of Pressure on the Positive Point-to-Plane Discharge in N_2, O_2, SO_2, SF_6, CCl_2F_2, A, He, and H_2, **H. C. Pollock, F. S. Cooper.** *Physical Review*, New York, N. Y., Vol. 56, 1939, pp. 170–175.

5. Gaseous Insulation for High-Voltage Apparatus, **G. Camilli, J. J. Chapman.** *AIEE Transactions*, vol. 66, 1947, pp. 1463–70.

6. Fluorine-Containing Gaseous Dielectrics, **G. Camilli, R. E. Plump.** *AIEE Transactions*, vol. 72, pt. I, May 1953, pp. 93–102.

7. The Dielectric Strength of Gaseous Fluorocarbons, **W. A. Wilson, J. H. Simons, T. J. Brice.** *Journal of Applied Physics*, New York, N. Y., vol. 21, 1950, pp. 203–05.

8. Improvements in or Relating to Electrical Apparatus Provided with Gaseous Insulating Atmospheres, **L. J. Berberich, C. N. Works.** British Patent *671780*, London, England, 1952.

9. The Mathematical Theory of Electrical Discharges in Gases, **Taro Kihara.** *Reviews of Modern Physics*, New York, N. Y., vol. 24, 1952, pp. 45–61.

10. Inelastic Collisions of Electrons in Helium and Townsend's Ionization Coefficient, **I. Abdelnabi, H. S. W. Massey.** *Proceedings*, Physical Society of London, London, England, vol. 66, sect. A, 1953, pp. 288–96.

11. A Condition on Uniform Field Breakdown in Electron-Attaching Gases, **Ronald Geballe, Marvin L. Reeves.** *Physical Review*, New York, N. Y., vol. 92, 1953, pp. 867–8.

12. Dielectric Materials and Applications (book), **A. von Hippel.** John Wiley and Sons, Inc., New York, N. Y., 1954.

13. Dielectric Breakdown of Sulfur Hexafluoride in Nonuniform Fields, **C. N. Works, T. W. Dakin.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 682–89.

14. Investigation of the Electrical Stability of Gases (in Russian), **B. M. Gokhberg, E. Ya. Zandberg.** *Zhurnal Teknicheskoi Fisiki*, Moscow, U.S.S.R., vol. 12, no. 2 and 3, 1942, pp. 65–75.

Discussion

A. W. W. Cameron (Hydro-Electric Power Commission of Ontario, Toronto, Ont., Canada): I am interested in the possibility of using some of the perfluorocarbon vapors, which the authors have so ably investigated, for location of gas leaks in high-voltage power cables of the nitrogen-pressure type. Presumably the usual type of halogen detector is sensitive to these vapors, and a leak could thereby be located if a small proportion of one of them were introduced into the nitrogen.

In cables so constructed that the nitrogen has access to the oil-paper dielectric, it appears to be inadvisable to introduce Freon or sulphur hexafluoride for this purpose, because cases have been reported in which these halogen gases had a deleterious action upon the oil-impregnated paper.

I hope that workers in this field will perform further research to determine the suitability of the perfluorocarbon vapors for this application. Besides the possibility of undesirable reaction with the dielectric, it would be advisable to consider the comparatively high boiling points of some of the vapors, with regard to the partial pressure which could be used in a cold cable. On the other hand, it might be found that some condensation would have no adverse effect.

Kenneth S. Wyatt (Phelps-Dodge Copper Products Corporation, New York, N.Y.): Power cables with oil-paper insulation have, always and inevitably, traces of moisture in them, and coaxial cables have polystyrene or polyethylene spacers which, like oil, give off small amounts of hydrogen under electric discharge. In such cables there are always sharp edges or local inhomogeneities which cause sharp local increases in electric stress, so that when higher frequency components are briefly superimposed on the normal 60-cycle wave, corona discharge occurs. It may be in only a small spot for a small fraction of a second, but the damage is done, hydrogen fluoride and hydrogen are formed, and as the process is cumulative, the cable will, after a time, be ruined. The hydrogen fluoride will, together with the moisture film on the cellulose fibers, greatly increase the dielectric loss. The hydrogen will markedly lower the ionization potential so that more corona discharge will occur. The rule therefore must be, never to introduce halogenated gases into paper-insulated cable where the electric stress may be sufficiently high to cause ionization.

By reading these papers, we gain the impression we are dealing with a completely halogenated compound such as SF_6, for example. Unfortunately, in practice there are always some unwanted variations or contaminations that are present at least in minute traces, such as SHF_5 or SH_2F_4, etc., some of which are unstable and most undesirable in oil paper insulation or even in coaxial cable with hydrocarbon spacers. In such cables, even if at first there is fairly pure halogenated gas, degradation takes place at high-stress points in the electric field. It might be practical to change the gas two or three times a year in coaxial cables, but this cannot be done in power cables.

It may be well to scrutinize more closely the foundations on which the authors of this and related papers base their studies and conclusions. To have enduring value, they should make clear how they purified the gas in each experiment, and state the moisture content and the number of bastard molecules per million, as these may sometimes have a marked bearing on the results, contrary to what some have said. These remarks are to give practical import to the theoretical and not to detract in any way from the excellent and indeed fascinating studies represented by this and allied papers.

G. Camilli, R. E. Plump, and T. W. Liao (General Electric Company, Pittsfield, Mass.): This paper is a major contribution to the state of our knowledge on gas dielectrics. It has special significance in the fact that liquids have been converted to gases systematically and their dielectric properties measured at comparable pressures. An extension of the number of substances on which information is now available has therefore resulted. Additional lines of work have also been opened. The authors should be complimented on the scope and thoroughness of the work which their data represent.

The results obtained with nitrogen are noteworthy because nitrogen is becoming more widely recognized as something more than an inert gas, good only for dilution. Fig. 7 clearly shows its beneficial dielectric effect in a highly divergent field. The fluorocarbon by itself suffers the anomalous loss of strength so common in electronegative gases under these circumstances. An explanation for the sustaining action of nitrogen can only be conjectural. One possibility might be that a disastrous accumulation of space charge is delayed by the presence of both the electropositive nitrogen and the electronegative fluorocarbon. The net effect might be one of mutual neutralization. Alternatively, especially under the stress of positive polarity, the nitrogen positive ions blanket the sharp edges of the electrode to the extent that the over-all field becomes less divergent and a higher gradient must therefore be applied to cause breakdown. Of course, the fluorocarbon can also produce a positive ion and an electron (the original phenomenon implied in the term "ionization potential"), but it appears that this process is less likely to occur when nitrogen is present. We are planning to study this possibility in the positive ion mass spectrograph.

In respect to pure SF_6, our results at atmospheric pressure in the uniform field of a Rogowski gap give a dielectric strength of 135 volts rms per mil, and between 3-inch disks with simply rounded edges the strength of SF_6 is 112 volts rms per mil. We should like to ask the authors what the actual breakdown distance was in their ozonizer cell and what this gap spacing yields as a comparable gradient with our results for SF_6. Also was either of the amines studied between the rod electrodes with varying pressure as was the case with C_7F_{14} in Fig. 7? In Fig. 6 the low strength under positive polarity impulse is definitely shown. What are the ideas leading to an explanation for such results? From the preceding comment it will be understood that we find only a few relatively minor details lacking. The work and the data are unreservedly impressive and informative.

J. J. Chapman (The Johns Hopkins University, Baltimore, Md.): This paper should provide further incentive for the study of the dielectric properties of complex gases. The detail given by the authors is commendable in an activity involving so many factors. Certain effects have been emphasized that are only slightly evident in simple gases. The effects of chemical structure on the breakdown strength of gases would appear to invite extensive research, with many compounds yet to be investigated. Although qualitative description remains because of many possible breakdown mechanisms, generalities are forming.

This work provides clues as to the effect of the resonating bond in C_7F_8 and also some comparison between straight chain and ring structures. The electronegative character of fluorine seems to have been reaffirmed as a factor of practical magnitude. It is encouraging to note that in spite of the vast number of compounds prompting investigation, classification and limitation are becoming evident. The elements, fluorine and chlorine to various degrees always remain in the discussion of high dielectric strength gases. The versatility of carbon chemistry keeps this element in the structural picture. It is interesting to note that in 1876, perfluors[1] of silicon and boron were mentioned as well as that of sulphur. Selenium, through SeF_6, has received attention in the exploration of gaseous dielectrics; see reference 14 of the paper.

The fact that C_7F_{14} mixed in relatively small proportions with N_2 yields a high dielectric strength medium at moderate pressure, appears useful. The ultimate explanation of such a phenomenon will be a matter of considerable interest.

It has been found that rather mild ultraviolet radiation decreases the dielectric strength of SF_6 to a quite noticeable extent.[2] Since no single property has been suggested as explaining breakdown results, further experiments with radiation may be relevant to this study of halogenated gases.

It would be of value to know the rate of solid deposition as a result of arcing in the C_7F_{14} mixture with N_2 as arbitrarily compared to some of the other pure halogenated methanes.

References

1. Sur la Formation et la Decomposition des Composes Binaires par l'Effluve Electrique, M. Berthelot. Comptes Rendus, Paris, France, vol. 82, 1876, pp. 1360–63. (In French)

2. The Electric Strength of Sulfur Hexafluoride at Radio Frequencies, J. W. Gibson, C. F. Miller. *Journal*, Electrochemical Society, Baltimore, Md., vol. 100, 1953, pp. 265–71.

L. J. Berberich, C. N. Works, E. W. Lindsay: One discusser, Mr. Cameron, suggests the possibility of using perfluorocarbon vapors for detecting the location of gas leaks in high-voltage power cables of the nitrogen-pressure type. As Mr. Cameron and Mr. Wyatt point out, the Freons and sulfur hexafluoride may have a deleterious action on oil-impregnated paper insulation. Pure perfluorocarbons are more stable than the Freons and therefore should have a much less deleterious effect on oil-impregnated insulation in the absence of corona. We are not primarily interested in high-voltage cable and did not investigate the effect of these perfluorocarbons on oil-impregnated paper. However, we did investigate the action of some perfluorocarbon compounds on several organic solid insulating materials, and found that satisfactory life can be expected in the absence of severe corona discharge. In the presence of corona, however, some breakdown of the perfluorocarbons will take place and the breakdown products will have an effect on impregnated paper insulation. Hence, in any impregnated-paper cable structure where corona is likely to occur the use of perfluorocarbons is not recommended.

In regard to Mr. Wyatt's question about the purity of these compounds, the perfluorocarbons were of the highest purity obtainable at the time the tests were made. Sufficient quantities of many of these compounds could not be obtained to permit any further treatment. In the case of C_7F_{14} and several other compounds, the material was redistilled and highly purified by chemical treatment. Such material is not degraded appreciably in the absence of corona. However, some degradation products can be expected in the presence of corona. These products may be adsorbed by such adsorbents as activated aluminum oxide and activated carbon.

The discussion by Mr. Camilli, Mr. Plump, and Mr. Liao of the results of mixtures of perfluorocarbon vapors with nitrogen is most interesting. These investigations[1] showed that fluorogases of lower molecular weight when mixed with nitrogen, also had a beneficial effect on dielectric strength. Thus, there can be little question that fluorocarbon gases, and vapors have a beneficial action when mixed with nitrogen especially under nonuniform field conditions. The breakdown mechanism of such mixtures is certainly not well understood at this time. We have not calculated the voltage gradients for our work with the ozonizer cell. The outside diameter of this cell is about $1^{5}/_{8}$ inches. The glass is about 5/64 inch thick, the air gap about 11/64 inch, and the inner glass about 1/16 inch thick. Since C_7F_{14} was measured at a fraction of an atmosphere in the ozonizer apparatus, the strength of SF_6 should be compared to that of C_7F_{14} as given in Fig. 1. For the uniform field produced by this sphere gap, the average voltage gradient is about 300 volts per mil rms for breakdown of C_7F_{14} at 1 atmosphere. Neither of the amines were studied between the rod gap of Fig. 7. In regard to a positive impulse breakdown lower than that for 60 cycles, as shown in Fig. 6, the breakdown mechanism of such mixtures is not well enough understood to attempt an explanation. However, the positive impulse strength of this mixture is much greater than that of nitrogen at the same pressure. Also, in a divergent field, the positive breakdown of most gases is usually much lower than the negative, as pointed out in the section on breakdown mechanism.

As Dr. Chapman points out, lack of complete information on the effects of chemical structure on the dielectric breakdown of gases indicates the need for further extensive research. In this paper, on the basis of the limited data, only a few general observations could be made. Many opportunities for further research are evident, and it is hoped that others will work in this interesting field.

Replying to Dr. Chapman's question on the rate of solid deposition as a result of arcing C_7F_{14}, it has been observed that about one half as much gas is formed by decomposition in arcing a fluorocarbon liquid as in the arcing of oil. It is believed that the amount of carbon deposited is proportional to the amount of gas evolved. Therefore, it is believed that much less carbon is deposited in the arcing of these perfluorocarbons than in the arcing of oil at the same arc energy. The rate of carbon deposition in oil circuit breakers is 0.65 gram per liter of gas generated by the arc. The amount of gas generated by the arc depends on the arc energy.[2] About 100 cubic centimeters of gas are evolved per kw-second arc energy for oil.

We wish to thank the discussers for their interest. The discussions form a valuable addition to the paper.

Reference

1. The Dielectric Behavior of Some Fluorogases and Their Mixtures, G. Camilli, T. W. Liao, R. E. Plump. *AIEE Transactions*, vol. 74, pt. I, Nov. 1955, pp. 637–42.

2. Reference 12 of the paper, p. 203.

Decicycle Magnetic Amplifier Systems for Servo Applications

LEO J. JOHNSON
ASSOCIATE MEMBER AIEE

S. E. RAUCH
NONMEMBER AIEE

Synopsis: The importance of minimizing time delays in a typical autopilot channel are discussed and the major advantages to be gained in the use of magnetic amplifier systems having responses which are small in comparison with the period of 1 complete cycle of the power supply are described. Magnetic frequency multipliers, magnetic pulse transformers, and transistor pulse generators are discussed, and corresponding circuit schematics which yield frequency multiplication from d-c and single-phase or multiphase a-c supplies are shown. A typical 400-cycle-per-second (cps) decicycle magnetic amplifier system which provides a reversible polarity d-c output is described. A comparison of response time, transfer characteristics, weight, and volume is made with a standard, high-performance 400-cps half-cycle response magnetic amplifier having the same power rating. It is shown that the decicycle magnetic amplifier system yields twice the sensitivity and one-sixth the response time for equivalent power rating and weight.

FROM the earlier self-saturating circuit, functioning as the basic magnetic amplifier element in servo systems, many circuit inventions and designs have resulted to increase the figures of merit and to decrease the response time of magnetic amplifiers. The advantageous properties of reliability and ruggedness to be found in the latter have been sufficient to motivate programs which have led to the many important advances. Magnetic amplifier circuits which once were limited to several cycles of delay have been improved by the availability of high-performance, rectangular core materials and by the introduction of circuits which yield a complete response within 1/2 cycle.[1,2] The adoption of pulse techniques,[3–4] which employ only magnetic and semiconductor materials, have made it possible to design decicycle magnetic amplifier systems which yield response times in the order of 1/10th to 1/100th cycle of the power supply frequency.[7] Response-time delays in servos can be reduced to negligible values, and inherent delays which produce undesired phase-lag characteristics in a system can be eliminated effectively.

In addition to the response-time improvements, the decicycle magnetic amplifier system can provide a reduction in physical size for a given power rating to values less than that required for equivalent amplifiers designed to operate at basic power frequencies of 60 or 400 cps. It has been shown[7] that the decicycle type of system can amplify reliably signals of a lower level than that employed in conventional circuits with equivalent response times.

In many cases, it can appear that the rapid response requirement for a magnetic amplifier in a closed-loop system is not of paramount importance; however a more serious analysis can show that the response characteristic of the amplifier is a limiting factor in determining the accuracy of the closed-loop system. As an illustrative example, consider a channel in a typical autopilot control such as is shown schematically in the block diagram of Fig. 1. In general, the dynamics of the aircraft require that the frequency spectrum be limited to values not exceeding 5 cps to provide smooth control of the aircraft and structural safety. In the case of a 400-cps system, a 1/2-cycle phase delay contributes a 2.25-degree (°) phase delay at the 5-cps frequency. Such a phase delay of 2.25° would not be expected to have any adverse effect in the 5-cps system provided it is a simple, single closed-loop system. Unfortunately, multiloop systems are usually encountered in present-day autopilots. Each control channel has a series of controlling functions, beginning with an input signal proportional to the desired course. A position gyro produces the signal reference corresponding to the actual heading of the aircraft. The difference between the desired course and gyro heading signal is amplified; the amplified signal then is operated upon by a stabilizing network, which, in turn, delivers a signal to a linear amplifier. The purpose of the stabilizing network is to compensate for dynamic characteristics of the aircraft to limit structural stresses in the airframe and to maintain smooth flight control with a minimum dynamic error. The output of the linear amplifier generally is fed into the control actuator such as an hydraulic valve, electric motor, or clutch. The actuator positions the control surface, and the control surface finally acts to change the position of the airframe with respect to the desired course.

To improve the dynamic response and accuracy of the flight control, it is useful to introduce a minor feedback loop within the afore-mentioned major loop. A position feedback from the control surface to the previously described linear amplifier insures the accuracy of the positioning of the control surface. In the case of the minor feedback loop, the linear amplifier gain requirement is at least an order of magnitude greater than that for the system without a minor feedback loop. By this means the control signal derived from the compensating network is accurately balanced to permit only a very small error between the actual signal from the compensating network and the position of the control surface. To provide stability in the minor loop and, at the same time, have no phase interference with the major loop, it is necessary to have a frequency response bandwidth for this minor loop which is at least an order of magnitude greater than the bandwidth of the major loop. When such a condition exists, it can be seen that a 1/2-cycle delay in the 400-cps magnetic amplifier can contribute a 25° phase delay. The 25° phase angle is no longer a negligible quantity, and in some cases can be suffi-

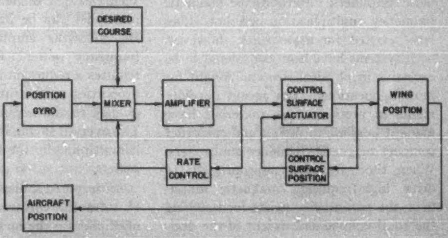


Fig. 1. Block diagram of a typical autopilot channel

Paper 55-542, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 10, 1955; made available for printing May 11, 1955.

LEO J. JOHNSON is with the Hufford Machine Works, Los Angeles, Calif., and S. E. RAUCH is with the University of California, Goleta, Calif.

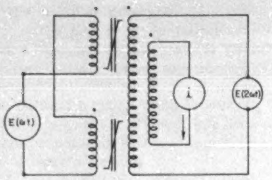


Fig. 2. Circuit schematic of a single-phase, magnetic frequency doubler

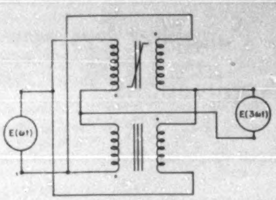


Fig. 3. Circuit schematic of a single-phase, magnetic frequency tripler

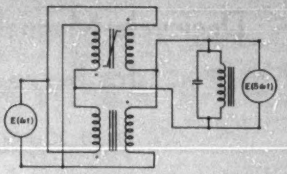


Fig. 4. Circuit schematic of a single-phase, magnetic frequency quintupler

cient to introduce instability in the minor loop.

From the foregoing it can be seen that employing the best 400-cps magnetic amplifier in such a system results in a marginal safety factor for maintaining stability. It appears desirable therefore to reduce the minimum phase delay of magnetic amplifiers to an insignificant value if it is possible. The decicycle amplifier has been developed for the specific purpose of making the phase delay negligible when operating in the most rapid closed-loop systems for present-day applications.

Decicycle Response Systems

The decicycle magnetic amplifier system employs wire-wound magnetic cores and semiconductors as the basic elements in the system. Both the saturable core and semiconductor components have been proved to be reliable and insensitive to shock vibration and temperature changes. Basically, a decicycle system produces a high-frequency power source from a lower frequency supply and uses the high-frequency power for operating a magnetic amplifier. As a result, the response characteristics of the system are limited to values which are a small fraction of 1 cycle of the low-frequency power supply.

The foremost problem which arises in the design of a decicycle system is the method of obtaining a reliable, efficient frequency multiplier which will produce sufficient power for the subsequent magnetic amplifiers. Methods for magnetic frequency multiplication without tubes have existed for some time; however, such systems have been considered to be excessive in physical size and weight for most applications. The recent developments in precision core materials have made it possible to design and construct compact magnetic frequency multipliers. When they are used appropriately to drive high-frequency magnetic amplifiers, the potentiality exists for reducing the total volume and weight of the decicycle systems to values less than that obtained by the more conventional magnetic amplifiers having the same rating and application. As the frequency increases, the physical volume and weight of the cores decrease approximately linearly for a given power rating. Since the amount of core material which is used in a stage of frequency multiplication is in the order of one-half that needed for an equivalently rated servo amplifier, it can be seen that, even with a low multiplication factor, a complete multiplier amplifier system will contain less core material than that which is required for a single servo amplifier operating from the basic low-frequency power source.

In addition to the reduction of magnetic core volume, there are many operational advantages which are available in using a decicycle magnetic amplifier system. For example, since the amount of magnetic material in the core of the servo amplifier can be reduced without any change in the power rating, the effect of a given signal will be greater. The relationship of response time versus control energy per unit core volume has been presented previously.[2] On the other hand, if response time is not to be made smaller than that which is to be found in conventional magnetic amplifiers, the signal sensitivity using a decicycle system can be increased. If a magnetic amplifier system using several amplifiers is employed, it can be shown that the savings in weight and volume are increased over the single amplifier. In such cases, a single magnetic frequency multiplier can be inserted to supply all the following amplifiers with a high-frequency power. Each amplifier contributes a reduction in size and weight in proportion to the ratio of the higher frequency to the original power frequency. The over-all saving in weight and volume is multiplied by the total number of amplifiers.

An important disadvantage of the use of saturable core reactors in frequency multipliers is the increased complexity of the multiplication system as the magnitude of the multiplication ratio increases. Such a system should be limited, therefore, to ratios less than 10. Frequency ratios greater than 10 generally require excessive amounts of materials and parts. As a result they no longer can offer a volume or weight advantage.

When decicycle systems are to be used with d-c amplifiers, the ripple frequency can be increased to a value such that it will not adversely affect the transfer functions of lead and lag networks commonly used in servo control loops. It should be noted that the low-frequency ripple in conventional magnetic amplifiers has introduced serious limitations for such circuits. Lead networks are used frequently to extend the linear frequency bandwidth of a servo system. If the lead network is to operate on a linear basis, the time constant of the network must be at least an order of magnitude greater than the period of the ripple frequency. In special cases, lead networks with time constants only three or four times the period of the ripple frequency are used, but in such cases the degenerative effect is apparent and must be compensated.

In special applications where very low-level signals are to be amplified, a saving in volume and weight can be achieved through the use of magnetic pulse generators which produce high peak pulses having the characteristics of very short time periods and long intervals of no voltage.[3] The use of spaced pulses for driving magnetic amplifiers provides the advantage of a considerable reduction in the volume of magnetic core for a given output level and, in addition, the number and size of wound core reactors for generating the power supply pulses from a low-frequency a-c line are small.

Some distinct disadvantages which arise, when using this system as described in reference 3 without modification, are the time delay between successive pulses and the maximum power output rating. Periods between output pulses often introduce excessive phase lags in a controlled system, and, therefore, the spaced

Vol.
Pt.
195

pulse technique cannot be used. The nature of the output pulses, i.e., being widely spaced in time, makes it very difficult to filter for a d-c output signal. If filtering is required, the response period of an adequate filter becomes long and the low efficiency in the filter results in a reduction of the total power gain of the system. The problem which arises when the output of the spaced pulse type of amplifier system is used to control a magnetic amplifier operating from a continuous source cannot be overlooked. The unfiltered pulses tend to produce equivalent spaced pulses in the output amplifier even when operating from a continuous source.

The use of newly developed semiconductor switches, such as switching transistors, provides another technique for the production of high-frequency power from either a low-frequency a-c or d-c source.[8,9] The advantages of using such a system are that very small volumes are required for the high-frequency section of the decicycle system and the number of components which produce the high-frequency in large frequency multiplication ratios is reduced. The transistor frequency multiplier offers a unique advantage of producing high-frequency power from a d-c source, which in some cases is an invaluable contribution to the designer of systems.

The temperature limitation of transistors at the time of this writing is a serious disadvantage, for the only transistor available for application is the germanium type. If and when silicon transistors are built for efficient switching applications, then the transistor frequency multiplier may be widely adopted for use in decicycle magnetic amplifier systems.

Types of Circuitry for Frequency Multiplication

MAGNETIC FREQUENCY MULTIPLICATION

Magnetic frequency multiplication can be accomplished with very good results for multiplication ratios in the range of 2 to 10. The basic methods of frequency multiplication which are adaptable to single-phase operation are the doubler, tripler, and quintupler. Combinations of the frequency multipliers can provide the additional multiplication values, except 7, in the range from 2 to 10.

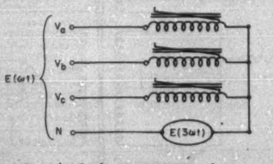


Fig. 5. A 3-phase, magnetic frequency tripler

SINGLE-PHASE DOUBLER

The circuit of a single-phase frequency doubler is shown in Fig. 2. Two windings in series, each winding on a separate saturable core, are connected to the fundamental frequency supply $E(\omega t)$. Two windings are wound common to the two saturable cores in such a manner that no voltage of fundamental frequency occurs across either winding while in operation. One of the common windings is supplied with a constant unidirectional current i. The other common winding is used to deliver to a load a double-frequency voltage $E(2\ \omega t)$. The two common windings can be simplified by using a single winding as a common carrier for both functions. The advantage of two windings is primarily for circuit isolation. Frequency doubling is obtained from the controlled separation of magnetization periods of the two cores. A more complete description of the mode of operation can be found in reference 10.

SINGLE-PHASE TRIPLER

The single-phase tripler, as shown in Fig. 3, employs a saturable reactor in series with a linear reactor. The saturable reactor winding is designed in general to absorb one-half of the supply voltage $E(\omega t)$. The linear reactor must absorb on an average basis the additional one-half of $E(\omega t)$. Frequency tripling arises as a result of the alternate storage and release of energy in the linear reactor. The saturable reactor performs as a gating device which determines the time distribution for tripling. The output wave form $E(3\ \omega t)$ should be filtered since this voltage has a large percentage of harmonics. Another characteristic of the circuit is that the output voltage $E(3\ \omega t)$, usually contains a component of the fundamental supply frequency, which is a subharmonic of the output. The relative proportion of subharmonic increases as the loading increases. Geyger[11] describes the operation of a single-phase tripler.

SINGLE-PHASE QUINTUPLER

The single-phase quintupler, Fig. 4, uses a circuit very similar to that shown for the tripler in Fig. 3. The major difference is the division of the supply voltage $E(\omega t)$ between the linear and saturable reactors. The saturable reactor is designed to saturate earlier in the half-cycle such that a maximum percentage of fifth harmonic is generated. A fifth-harmonic energy-storage resonant circuit, which is shown at the output terminals of the voltage $E(5\ \omega t)$, is recommended to increase the percentage of fifth multiple as well as to control the wave form. The energy-storage filter should be designed to be able to store approximately 10 times the energy delivered to the load on the basis of a single cycle of operation.

MULTIPHASE MULTIPLIERS

Multiphase frequency multipliers differ fundamentally from single-phase units in that there occurs essentially no energy storage and that only one saturable reactor per phase is needed. Multiplication values in multiphase multipliers are obtained by the addition through time phasing of individual phase pulses across a common output. It may be noted that, although the supply is multiphase, the

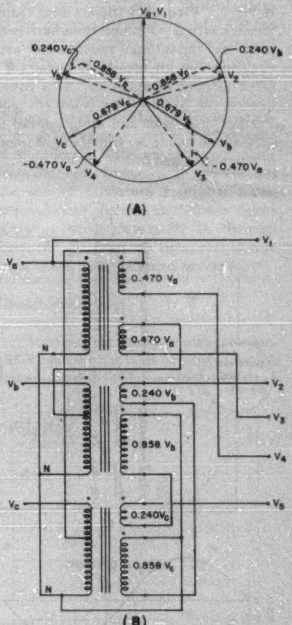


Fig. 6. A 3-phase, magnetic frequency quintupler

A—Phase transformation from a 3-phase supply

B—Circuit schematic showing electrical connections

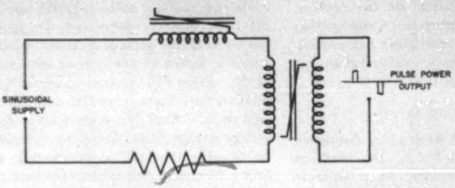


Fig. 7. A simple magnetic pulse transformer system

output is single phase for this type of system. Multiphase multipliers produce odd integer multiplication ratios.

The 3-phase frequency tripler differs from the single-phase tripler in that only one saturable reactor per phase without any linear reactor is used. The frequency multiplication results from the phase delay in saturation of each reactor in the separate supply phases, and the multiplication ratio is directly proportional to the number of phases used. A schematic diagram of an efficient frequency tripler which employs 3-phase excitation is shown in Fig. 5. Frequency triplers using transformers with $\Delta - Y$ connections have been built, but the internal resistance of such configurations are inherently higher for a given amount of material. As a consequence, the $\Delta - Y$ system is less efficient. Further discussion and design detail of the 3-phase frequency tripler can be found in reference 5 and 6.

Since, in general, only two or three multiphase supplies are available, it becomes necessary to introduce phase-changing transformers to obtain the required spaced phase voltages for the higher frequency multiplication ratios. Fig. 6(A) presents a phasor diagram showing the method by which phase transformation is accomplished for a quintupler with a 3-phase supply. Fig. 6(B) is a schematic diagram showing the electrical connections for the corresponding phasor diagram of Fig. 6(A). This method can be easily generalized so that any odd frequency multiple can be obtained. A quintupler which has been built according to this technique has been reported previously in reference 7.

MAGNETIC PULSE TRANSFORMERS

A schematic diagram of a simple magnetic pulse transformer system is provided in Fig. 7. The circuit uses two saturable reactors and a resistor in series with a sinusoidal power supply. One of the saturable reactors magnetizes at a lower magnetizing force than that required for the other reactor. Its function in the circuit is to absorb voltage from the sine wave supply during the early period of each half-cycle. When the first saturable reactor reaches saturation, the second reactor starts to absorb voltage and transforms this voltage to a secondary winding until its core saturates. In this manner, there results a pulse output having a negligible rise and decay time and a near rectangular wave characteristic. The first series saturable reactor isolates the low-magnitude portion of the sinusoidal supply from the output. The second saturable reactor transforms a small portion of the supply into a short-duration square-wave output, and finally the resistor limits the supply current during the remaining portion of each half-cycle when both cores are saturated.

To reduce the relaxation time to a minimum and to improve the efficiency of the pulse transformation system, a general circuitry, Fig. 8(A) can be employed. This system can produce a continuous pulsed output in contrast to the intermittent power which arises from circuits such as described in Fig. 7. Fig. 8(A) requires only saturable reactors for use as transformers and, therefore, makes the entire voltage wave useful at the terminals of the output. The time duration in magnitude of the output pulses can be adjusted by the winding constant for each saturable reactor. A bias-current source determines the proper firing interval for each reactor and thereby controls the pulse output continuity and wave form. The wave diagram of Fig. 8(B) indicates a typical input-output voltage characteristic when all saturable transformers are identical except for the bias windings. Fig. 8(B) is only illustrative and is not to be considered as an optimum characteristic. The adjustment of the winding constant will enable the designer to obtain a constant amplitude for all pulses in the output.

The output current is limited by the bias current, and therefore the maximum

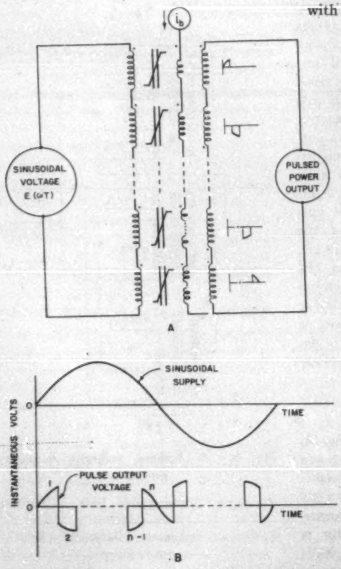


Fig. 8 (left). (A) A circuit schematic for a continuous pulse transformation system. (B) Typical output voltage characteristic for a sinusoidal supply

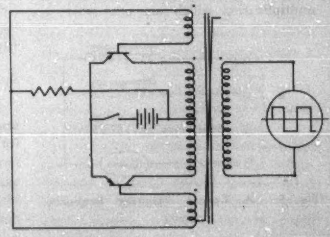


Fig. 9 (right). Circuit schematic of a transistor pulse generator

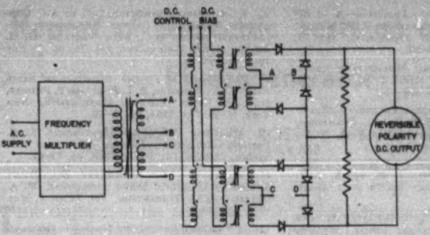


Fig. 10 (left). Circuit schematic for a sample decicycle magnetic amplifier system which delivers a reversible polarity d-c output

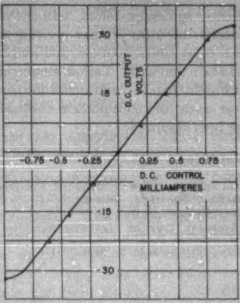


Fig. 11 (right). Transfer characteristic of the sample decicycle system

output current is directly dependent upon the bias when the saturable cores are of the same configuration and materials. External bias is not necessary if core and winding configurations are adjusted individually. For large frequency multiples, the individual core configuration control becomes cumbersome, and therefore the bias-current technique should be favored. It should be added that the pulse frequency multiplier is operated from a single-phase supply and does not require phase-changing transformers. Magnetic pulse transformers should be given serious consideration when only single-phase power is available.

Transistor Pulse Generator

The schematic of Fig. 9 shows the circuit for a transistor pulse generator. A direct-voltage supply, two transistors, a saturable transformer, and a resistor are the basic elements in the circuit. To obtain the best results from this circuit, a magnetic core having a very rectangular hysteresis loop should be selected. Such a core will then switch the transistors on and off rapidly and will reduce the switching power dissipation. The output frequency then is dependent upon the d-c supply voltage and the winding constants. A more complete analysis of this circuit can be found in the literature.[8,12]

Several disadvantages are inherent in the transistor pulse generator. The power rating is limited by the transistor power rating, which at the present time is relatively small. Germanium transistors' deratings caused by high temperatures limit the practical application of transistor multipliers to ambients not exceeding 85° centigrade. The frequency of the output of transistor multipliers of the type shown in Fig. 9 can be regulated only by controlling the supply voltage and as a consequence, frequency drift can become objectionable when unregulated d-c power is used. The use of separate switching transformers will provide controlled transistor switching and result in a controlled output frequency regardless of the variation in d-c supply voltage.

Table I. Comparative Data and Characteristics of a Conventional, High-Performance, Half-Cycle Response Magnetic Amplifier and the Decicycle Response Magnetic Amplifier System

	High-Performance Half-Cycle Response Magnetic Amplifier	Decicycle Magnetic Amplifier System
Supply frequency, cps	400	400
Single-phase supply voltage, volts	115	115
Maximum power output, watts	1	1
Duty cycle	Continuous	Continuous
Control power required for full output, microwatts	120	60
Minimum response time for full change in output, milliseconds	1.2	0.2
Minimum signal sensitivity, microwatts	2	1
Linearity, per cent	5	5
Maximum power gain	8,000	16,000
Control frequency for 45° phase shift between control signal and output, cps	50	300
Total volume, inches^3	7.0	7.5
Total weight, ounces	11	12

Typical Decicycle Magnetic Amplifier System

A decicycle magnetic amplifier system employs a frequency multiplier to increase the basic supply frequency to a higher multiple and a high-performance magnetic amplifier operating from the carrier frequency delivered by the multiplier. A choice of the frequency multiplier can be made from the types discussed in the foregoing, the particular selection being governed by the available power source and the desired output characteristics. In accordance with the selection of the frequency multiplier, the magnetic amplifier design is based upon the multiplier's output frequency and the degree of performance of the end use.

Fig. 10 represents a decicycle magnetic amplifier system which can produce reversible polarity d-c output. Such an amplifier has application in controlling hydraulic servo valves, d-c servo motors, and other similar devices. The transfer characteristic which shows the performance of a sample decicycle system is given in Fig. 11. Fig. 12 illustrates the transfer characteristic of a comparative 400-cps system. Table I presents data on the general characteristics of a 400-cps system in comparison with the equivalent decicycle system. From this table, it can be seen that the sensitivity of the decicycle system is improved by a factor of 2, and the response time is decreased sixfold. The improvement in performance of a decicycle system is understandable since the higher frequency supplies permit the use of miniature saturable cores, which, in turn, require much lower values of energy for complete reset. A thorough discussion of response time versus core volume has been presented previously.[2]

Conclusions

A summary on decicycle magnetic amplifier systems has been presented, which find application in high-performance servo

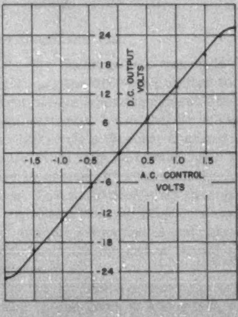


Fig. 12. Transfer characteristic of a comparative 400-cps magnetic amplifier

systems, automatic pilots, telemetering devices, and other similar systems for which delay times, volumes, and weights are to be minimized. Various techniques which yield frequency multiplication from d-c or single and multiphase a-c supplies have been discussed.

References

1. On the Mechanics of Magnetic Amplifier Operation, **Robert A. Ramey.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

2. Response Time of Magnetic Amplifiers, **Leo J. Johnson, S. E. Rauch.** *Electrical Manufacturing*, New York, N. Y., Sept. 1954.

3. Pulse Relaxation Amplifier—A Low-Level D-C Magnetic Amplifier, **R. E. Morgan, J. B. McFerran.** *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 345–50.

4. Magnetic Frequency Multipliers, **Leo J. Johnson, S. E. Rauch.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 448–52.

5. Odd Integer Magnetic Frequency Multipliers, **Leo J. Johnson, S. E. Rauch.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 43, 1954, pp. 168–74.

6. Analysis and Design of a Magnetic Frequency Multiplier, **Otto J. M. Smith, J. T. Salihi.** *AIEE Transactions*, vol. 74, pt. I, March 1955, pp. 99–106.

7. Single Phase Magnetic Amplifier Systems With Decicycle Response, **Leo J. Johnson, S. E. Rauch.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 10, 1954, pp. 162–67.

8. A Switching Transistor D-C to A-C Converter Having an Output Frequency Proportional to the D-C Input Voltage, **G. H. Royer.** *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 322–26.

9. Transistors as On-Off Switches in Saturable-Core Circuits, **R. L. Bright, G. F. Pittman, Jr., G. H. Royer.** *Electrical Manufacturing*, New York, N. Y., Dec. 1954, pp. 79–82.

10. Magnetic Frequency Conversion, **L. C. Harriott.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 9, 1953, pp. 78–87.

11. Elimination of Asymmetry Zero-Drift Errors in Magnetic Servo Amplifiers, **W. A. Geyger.** *AIEE Transactions*, vol. 74, pt. I, May 1955, pp. 239–45.

12. A Variable Frequency Magnetic-Coupled Multivibrator, **R. L. Van Allen.** *AIEE Transactions*, vol. 74, pt. I, July 1955, pp. 356–61.

◆

Discussion

R. O. Decker (Westinghouse Electric Corporation, Pittsburgh, Pa.): The authors have made a rather extensive survey of frequency multiplication but a few points pertaining to the transistor pulse generator need clarification.

First, with regard to core material, a magnetic core with a very rectangular hysteresis loop is not essential for square-wave switching. Ferrite cores, as well as C-cores made of Hipersil, have been used to produce square-wave outputs.

Second, with regard to temperature considerations, it is true that the transistors must be derated. However, the 65 C limit is too low. A practical supply, which uses two commercially available 20-watt transistors, has been built and tested. This power supply delivers 8 watts continuously at 71 C, at a frequency of 10 kc. Other supplies have operated satisfactorily at 85 C.

Third, with regard to the frequency drift, it should be kept in mind that the square-wave oscillator with one core has a constant volt-second area output. Since the magnetic amplifier is primarily sensitive to volt-second area, frequency changes have small effect on amplifiers operated from the transistor power supply. For example, if the magnetic amplifier cores are designed to absorb the volt-second output of the square-wave oscillator, the cores can always be controlled over their full range, even with drastic frequency changes.

Report of Dielectric Tests on a Large Hydrogenerator

C. A. DUKE
MEMBER AIEE

C. W. ROSS
ASSOCIATE MEMBER AIEE

J. S. JOHNSON
MEMBER AIEE

THE OBJECT of the tests was to gather information regarding the characteristics of the insulation of a large generator winding. The tests also provided an opportunity to evaluate new test equipment and test methods. In particular, it was desired to evalute certain factors which may influence the results of various dielectric tests.

The machine tested was a 40,000-kva, 13.8-kv hydrogenerator which had been in service for 12 years. Since the stator was to be rewound the opportunity was presented of making destructive tests on the old winding prior to the rewind.

The various types of tests made were as follows:

1. Dielectric-absorption tests versus temperature and grounding time.
2. Power factor and capacitance tests versus temperature.
3. Slot discharge test.
4. High-voltage d-c tests to predict insulation failure.
5. A-c and d-c breakdown tests.

Discussion

Dielectric-Absorption Versus Temperature Test

The dielectric-absorption-versus-temperature test consisted of a series of dielectric absorption tests made with a 2,500-volt cross-coil ohmmeter with motor-generator supply and was begun as soon as the generator was cleared from the line. The tests were continued, alternately making absorption tests and then discharging the winding, until the generator temperature cooled to 30 degrees centigrade (C). The purpose of the test was to evaluate the effect of the temperature of the winding on the dielectric absorption curves.

The dielectric-absorption curves measured at 5 C intervals are shown in Fig. 1. The resistance curve versus time measured at 45 C shows a bend upward after about 8 minutes and indicates a relaxation time of the absorption current.[1] As the temperature decreases the relaxation time is increased and is seen from Fig. 1 to be greater than 15 minutes for temperature below 40 C. Previous tests, made on June 5, 1947, and shown in Fig. 2, indicate that the relaxation time occurs at less than 1 minute at 68 C. The effect of such a relaxation time for the absorption current is to increase the measured value of the polarization index, PI, as shown in Table I. For measurements made at time appreciably less than the relaxation time and at lower temperatures, the PI is essentially invariant with temperature, even though the resistance at any given time varies considerably with temperature.

The apparent resistance values measured at 1, 10, and 15 minutes are plotted versus temperature in Fig. 3 for the purpose of evaluating the temperature coefficient A for this particular insulation.[2]

$$A=\frac{\log_{10}\dfrac{R_{T_1}}{R_{T_2}}}{T_1-T_2}$$

where

R_{T_1} = resistance at temperature T_1
R_{T_2} = resistance at temperature T_2

The A value is 0.017 for the 15-minute tests, 0.0217 for the 10-minute tests, and 0.0263 for the 1-minute tests. Although the temperature range was not broad enough for a general evaluation, the results point out possible variations of the correction factor. Variations in this case are due in part to the relaxation time noted for the absorption current at higher temperatures.

Power Factor and Capacitance Tests

A 60-cycle-per-second capacitance bridge was utilized to measure the winding capacitance and power factor at intervals as the generator cooled from 45 to 30 C. The data are shown in Table I. Power factor shows a negligible change with temperature and capacitance increases slightly as temperature decreases. The capacitance showed a small increase after the machine had remained at a room temperature of 27 C for about 10 days, which may have been because of moisture being taken into the winding.

Dielectric Absorption Tests with Varied Grounding Time

The dielectric-absorption tests were performed about 10 days after those just discussed. The purpose was to demonstrate the effect of insufficient discharge time on the shape of the dielectric-absorption curves. The test steps and results are given in Fig. 4. The first 2,500-volt dielectric-absorption curve is also shown in Fig. 4 for comparison with previous data. This curve has a decreasing slope for longer charging times, probably due to moisture which increased the leakage current and resulted in a lower PI.

The effect of insufficient discharge on the first 500-volt dielectric-absorption curve is to give an apparent resistance versus time and a PI which are too large. Since the first applied voltage was 2,500 volts and sufficient discharge time was not used before applying the 500-volt step, there is an effective negative-absorption current for time greater than about 1.5 minutes. This negative absorption current causes the measured resistance to swing higher than the actual leakage value, then decrease slowly toward the leakage resistance. The second 500-volt dielectric-absorption curve approaches more nearly the correct values and would indicate moisture since it lies above the 2,500-volt curve. The effect of insufficient grounding time would have been greatly reduced if the lower voltage had been applied first. The last 2,500-volt curve lies essentially on the first and shows the reproducibility of the tests.

Slot Discharge Test

The slot discharge test was made by determining the 2,000-cycle voltage component present when the generator insulation is stressed by a 60-cycle voltage. The 2,000-cycle component is caused by an arc which may be formed when the coil sides are not grounded to the laminations. The arc formed when the capacitance voltage built up between the coil side and the lamination is sufficient to break down the air gap is termed slot discharge.[3]

For the slot discharge tests a new setup similar to that given in reference 3, was used. The generator insulation (equivalent to a capacitor) was stressed by using

---

Paper 55-236, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted November 30, 1954; made available for printing April 20, 1955.

C. A. Duke is with the Tennessee Valley Authority, Chattanooga, Tenn.; C. W. Ross is with the James G. Biddle Company, Philadelphia, Pa.; and J. S. Johnson is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

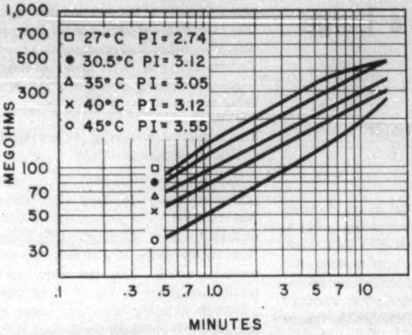


Fig. 1. Dielectric absorption tests at various stator temperatures

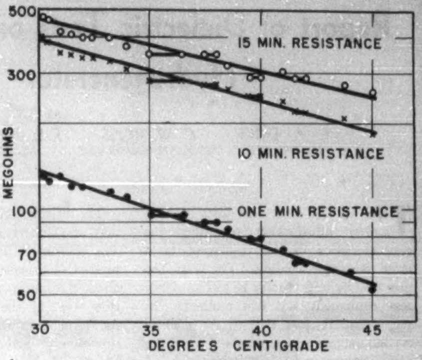


Fig. 3. Resistance versus temperature

the resonant rise of voltage in a series circuit which is near resonance. Four reactors were used in the series circuit. Each reactor consisted of a 2,300-volt winding of a 5-kva distribution transformer, with a 1/4-inch gap in the magnetic circuit. Then with an 800-volt supply, the voltage across the generator insulation was 8,000 volts as measured by a voltmeter built into the slot discharge test set. The 800-volt supply was obtained from a 5-kva 440/2,200-volt test transformer supplied from a 0/220-volt 7.5-kva variable voltage autotransformer.

A circuit is built into the slot discharge test set to simulate the arcing of one coil side. This arc produces a 2,000-cycle signal in the test circuit. The magnitude of the signal when displayed on a cathode-ray oscillograph screen is used as a reference for determining the relative intensities of the slot discharge. For calibration the test circuit is shunted by a fixed capacitor equal to the capacitance of the winding tested. The reference calibration used in this test to simulate one coil side arcing gave a deflection of 1/4 inch on the oscilloscope screen for 8 kv across a 0.5-microfarad capacitor. The result of the slot discharge tests are listed in Table II.

Each phase tested showed definite evidence of slot discharges with phase *A* showing more than phase *B*, which in turn was greater than phase *C*. The discharges could be detected at any voltage over 5 kv, phase to ground. The magnitude of the signal increased at a more rapid rate than the line-to-ground voltage. No good estimate of the number of coil sides which were arcing could be made since the total discharge energy varied with the length of the arc gap and the resistance of the conductive coating. From the over-all readings, however, it can be stated that at least three coil sides in each phase were arcing.

Examination of the coils after their removal showed considerable slot discharge damage and in some cases the damage was quite severe. All of the insulation failures occurring during subsequent a-c and d-c tests were later found to have occurred at points of severe discharge damage.

High-Voltage D-C Tests to Predict Insulation Failure

There is a published theory that the d-c breakdown voltage of a winding may be predicted by measuring the winding insulation resistance at lower d-c voltages and extrapolating the resistance-voltage curve to breakdown.[4-6] Thus, the insulation strength of a winding could be predicted without applying enough direct voltage to cause actual insulation failure.

A number of attempts were made to follow this procedure and predict the insulation failure voltage. The effect of varying the charging time for each step voltage and the polarity of the test voltage was also investigated. The test procedure and results of the various tests are described in the following.

Table I. Capacitance, Power Factor, PI, and Resistance versus Stator Temperature

Stator Temperature, C	Capacitance		PI	Resistance		
	Microfarads	Per Cent		1 Minute	10 Minutes	15 Minutes
45			3.5	52.5	185	260
43	1.26	5.2	3.5	60	210	275
42	1.27	5.25	3.46	65	220	290
41	1.27	5.25	3.34	72	240	305
40	1.27	5.25	3.16	79	250	290
38.5	1.28	5.25	3.12	85	265	320
38	1.28	5.3	3.05	90	275	350
36.5	1.29	5.3	3.0	95	285	350
35.5	1.3	5.2	3.0	100	300	350
34.2	1.28	5.3	2.93	105	310	375
34	1.29	5.3	2.83	110	310	375
33.2	1.29	5.3	2.83	115	330	400
32.5	1.29	5.3	2.83	120	340	400
32.1	1.29	5.3	2.83	120	340	400
31.5	1.30	5.2	2.92	120	350	410
31	1.31	5.2	2.72	130	355	420
30.5	1.32	5.5	3.12	125	390	460
30.2	1.31	5.2	3.06	130	400	475
After ten days						
27	1.36	5.2	2.74	150	410	440

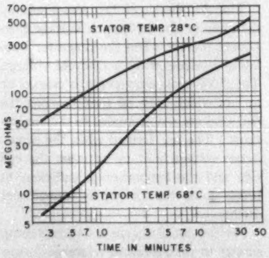


Fig. 2. Dielectric absorption test made June 2–5, 1947

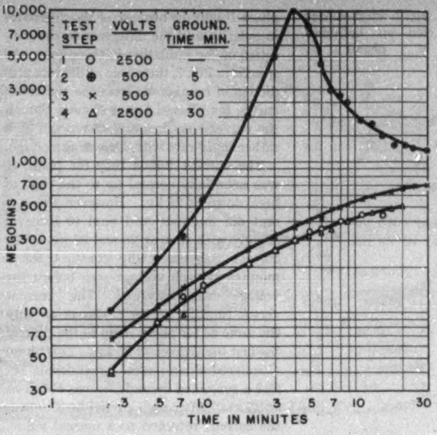


Fig. 4. Dielectric absorption test with various grounding times

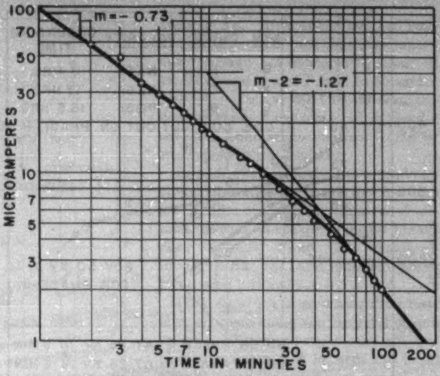


Fig. 6. Discharge current versus time after first high-voltage test

All high-voltage d-c tests involving sensitive current measurements up to 40 kv were made with a commercial dielectric test set of new design. Facilities for reversing the polarity of the test voltage, and making current and voltage measurements of both polarities were available in this equipment.

Short-Time High-Voltage D-C Test, Positive Polarity

The positive-polarity test procedure was as follows: with negative polarity to ground, voltage was applied for 10 minutes per 5-kv voltage step and the current versus time was recorded for each voltage step.

The results of the calculated resistances at the end of 10 minutes for each voltage step are plotted in Fig. 5. The resistance-voltage curve shows a slight drooping characteristic with increasing voltage. This test was stopped at 20 kv when an increase in current versus time was noted at 20 kv after about 7 minutes. Fluctuations of the current meter were also noted when the current tended to increase and were taken as indicating possible failure.

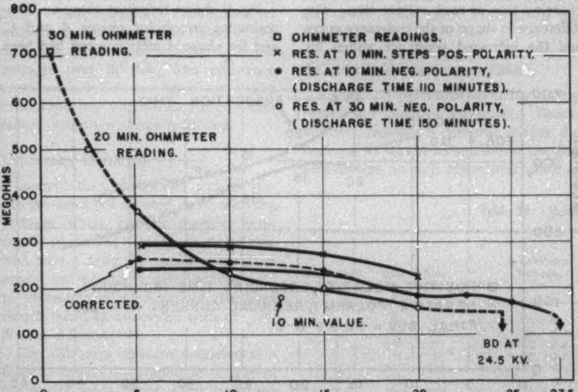


Fig. 5. High-voltage test on complete stator winding

So that additional tests could be made on the winding in its initial condition the voltage was removed at the end of the 20-kv step and measurements of current versus time were made as the winding was being discharged. The discharge current versus time is plotted in Fig. 6. The use of this curve for determining the equation for the absorption component of current may be somewhat in error since the generator winding was not fully charged before starting the discharge measurements.

Short-Time High-Voltage D-C Test, Negative Polarity

The negative-polarity tests procedure was similar to that just given except that the positive polarity was connected to ground. The results of the calculated resistances at the end of 10 minutes for each voltage step are plotted in Fig. 5 for comparison with the results of the positive-polarity test. The corrected curve for the previous charge is shown by the dashed curve. The resistance curves have a similar shape up to 20 kv. The difference in resistance values of the two curves may be due to the difference in polarity of the applied voltage. According to the Evershed theory this would indicate moisture in the winding.[7]

The 10-minute resistance values for

Table II. Result of Slot Discharge Tests

Voltage, Kv	Deflection on Screen, Inches		
	Phase A	Phase B	Phase C
4.5			
5	3/4–1¼	1/8–3/16	1/8–3/16
6	3/4–1¼	3/8–3/4	3/8–1/2
8	1–2¼	1¼–2	5/8–1

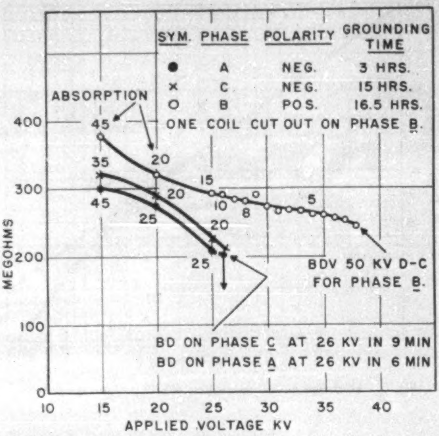


Fig. 7. High-voltage test on individual phases

voltages over 20 kv are higher than would be expected from the general trend of the resistance curves. This apparent leveling off of the resistance is due in part to the fact that smaller voltage steps were taken and the absorption current at the end of 10 minutes is proportionally decreased. This test was stopped after about 2.5 minutes at 27.5 kv when the current was noted as increasing at a progressively more rapid rate with time.

Long-Time Test, Negative Polarity

The purpose of this test was to show the effect of longer charging times on the resistance-versus-voltage curves. The values of the final calculated resistance values versus voltage are plotted in Fig. 5 for comparison with curves from the short-time tests.

The resistance decreases rapidly with increased voltage in the lower voltage region and the curve falls in line with the previous 2,500-volt ohmmeter readings. The 12.5-kv resistance value falls below the curve because of a shorter charging time of 10 minutes. The 15-kv 30-minute value is slightly above the curve because of the intermediate 12.5-kv voltage step. The nearly flat resistance curve for voltage below 15 kv on the short-time readings is apparently due to the compensating effect of leakage and absorption currents. As more voltage steps are taken the short-time resistance curve should tend to approach the long-time curve, the short-time resistance values being the lower. The fact that the long-time resistance values are lower in this case indicates that some damage was done to the insulation by the previous test at the 27.5-kv step. That a weak spot in the insulation may have been deteriorated by the previous test was substantiated by the fact that breakdown occurred suddenly at 24.5 kv as the applied voltage was being increased to 25 kv.

High-Voltage D-C Test on Individual Phases

Coil 25 of phase *B*, which failed in the previous (full-winding) tests was cut out and the testing was continued on the individual phases. The curves of resistance versus voltage obtained for each phase tested are given in Figs. 7 through 9.

The resistance curves in Fig. 7 for the three phases have essentially the same charging time at each voltage step. The difference in shape of the resistance curve and the increased resistance values for phase *B* may be because of the difference in polarity and the difference in the condition of the insulation in this phase. As seen from Fig. 7, there was no definite prediction of breakdown from the resistance curves for phases *A* and *C* even though the curves were bending downward at a rather rapid rate with increasing voltage.

The current versus time for phase *A* was essentially normal up to the time of breakdown at 26 kv. However, the current did fluctuate and tend to increase slightly with time at 25 and 26 kv. The current was essentially constant for 5 minutes at each voltage step before the voltage was increased. The current versus time for phase *C* was essentially the same as for phase *A* up to the time of the first breakdown at 25 kv. This phase *C* suffered three partial breakdowns before a sustained breakdown was obtained at 26 kv. After each partial breakdown, the current returned to a normal value. The breakdown in nearly all cases took place some time after the breakdown voltage was applied. This would indicate some deterioration of a weak spot in the insulation with time.

The two resistance curves shown in Fig. 8 for phase *B* with positive and negative polarities are not directly comparable since the charging times at the various voltage steps are different. The residual current and the shorter charging times on the negative-polarity test will tend to make the resistance values lower than those obtained during the positive-polarity test. These factors in the measurement make the resistance curve appear to have a hump or maximum near 20 kv.

Fig. 9 shows the short charging time-resistance curves for phases *A* and *C*, and for phase *A* after one coil had been

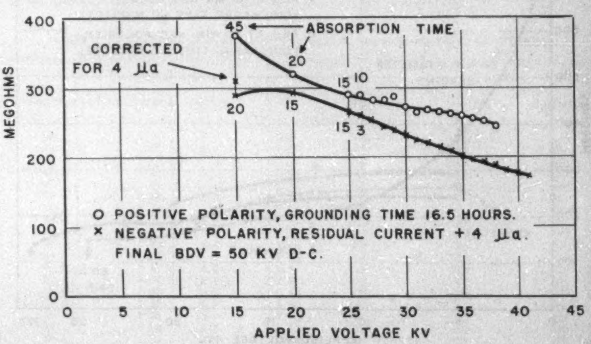


Fig. 8. Test on phase B, one coil cut out

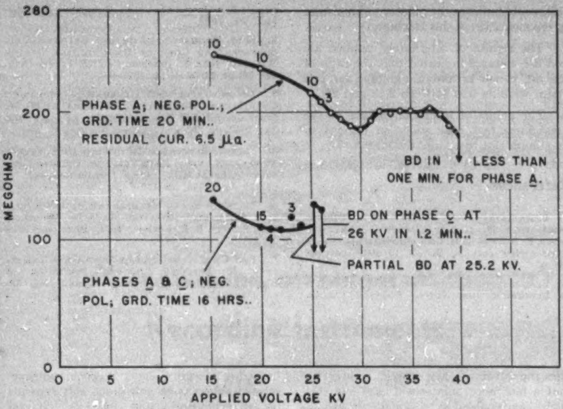


Fig. 9. Test on phases A and C, and on phase A, one coil cut out

cut out of each phase. The curve for phases *A* and *C* does not indicate a pending failure near the breakdown voltage of 25.2 kv. However, the current-time values did show erratic behavior prior to the partial breakdown. Discharges were also audible before the partial breakdown. After the partial breakdown the resistance increased; however, the current appeared to be increasing slowly with time. Increasing the voltage to 26 kv did not cause an abnormal current change, and breakdown occurred without warning in about 1.2 minutes.

The resistance curve for phase *A*, Fig. 9 showed an abnormal increase after about 30 kv. This could have been associated with the rather large pulses noted on the discharge detector (to be described) for voltages over 30 kv. The resistance curve tends to decrease again just before breakdown but not to such a degree that failure would have been predicted.

DISCHARGE DETECTOR

Current fluctuations and audible discharges were often noted at or near the voltages where current readings were erratic or showed a tendency to increase with time. After the test of phase *A* an improvised discharge detector was used to determine whether or not discharge signals could be associated with the prediction of breakdown.

The discharge detector consisted of a filter to reduce the amplitude of current harmonics below about 2,000 cycles per second. The output of the filter was connected to a cathode-ray oscilloscope and the input signal was obtained from a 1,000-ohm resistor in series with the ground return to the high-voltage test set.

No correlation was obtained between current fluctuation and signals on the cathode-ray oscilloscope, which may have been owing to the lack of detector sensitivity or to the difficulty of observing visually the random pulses.

In the cases where discharges were noted, either audible or on the cathode-ray oscilloscope, their cause may not have been directly related to deterioration or pending failure of the insulation. The discharges may have been taking place either at the points where coils were cut out or at some point external to the winding.

A-C AND D-C BREAKDOWN TESTS

A summary of failures occurring during testing are given in Table III. Breakdowns 1 through 6 occurred during d-c testing. It should be noted that the first breakdowns on each phase took place at about 25 kv. After one or two bad coils had been cut out of each phase, the breakdown voltage was increased to about 40 kv for phases *A* and *B*, and to 50 kv for phase *C*. This would indicate that the damage versus time suggested in the d-c tests was confined to incipient faults or weak spots in the insulation.

Breakdowns 7 through 11 were made by alternately applying 1-minute d-c and a-c voltages of increasing values to determine the ratio of the d-c to the a-c breakdown voltage. From the breakdown tests 7 and 11 the ratio indicated would be between 2.12 and 2.3.

The coils which failed during these tests were removed and given a careful examination. Wherever possible, the point of insulation failure was located, and in all cases this was at a point of maximum slot discharge damage.

Summary of Results

1. The absorption current component may have a relaxation time which at higher temperature may be of the order of the dielectric-absorption test time. This results in resistance values for longer times and PI values which are too high.
2. The temperature correction factors for dielectric-absorption curves may vary somewhat for the various times of electrification at which the resistance is determined.
3. The effect of insufficient discharge time on the 2-voltage dielectric-absorption measurement was demonstrated. This effect is to give resistance-time values which are too high. The effect is more pronounced when the first applied voltage step is larger than the second. In this case it is possible to obtain resistance values which are larger than the true leakage resistance. The PI measured with insufficient discharge time will also be high.
4. The curves of resistance versus voltage were shown to be affected by the length of charging time at each voltage step. The effect of insufficient charge time is to give resistance values which are too low. This effect is more pronounced in the lower voltage region of the curves where the absorption current is relatively large compared to the leakage currents.

Table III. Summary of Coil Failures

Breakdown No.	Coil No.	Slot No.	Phase	Breakdown Voltage, Kv	Time to Breakdown	Prebreakdown Voltage, Kv	Ratio D-C Breakdown Voltage / A-C Breakdown Voltage
1	25	25	B	−24.5 d-c	0		
2	150	150	A	−26 d-c	7 minutes		
3	238	238	C	−26 d-c	10 minutes		
4	28	28	C	−26 d-c	1 minute		
5	346	346	A	−40 d-c	> 1 minute		
6	110	110	A	−39 d-c	0	−40 d-c	
7	358	358	B	19.3 a-c	0	−41 d-c	> 2.12
8	376	376	C	23.6 a-c	22 seconds	−45 d-c	> 1.9
9			A	−39 d-c	0	−39 d-c	
10			B	21.7 a-c	20 seconds	−45 d-c	> 2.07
11	237	242	C	−50 d-c	33 seconds	21.7 a-c	< 2.3

5. It was not possible to predict the breakdown voltage from any of the resistance-voltage curves made on the complete stator winding or on the individual phases. In some cases it appeared that erratic current behavior with time and discharges noted at voltages near breakdown may have been associated with the pending failure.

6. Many of the breakdowns which occurred during the d-c tests took place some time after the breakdown voltage had been applied. This would indicate some deterioration of a weak spot in the insulation with time application of direct voltage.

7. The results of the polarity tests indicate that resistance values obtained with negative polarity to the winding are lower than those obtained with positive polarity. According to the electro-endosmosis theory this would indicate the presence of moisture in the insulation of this machine.

8. The results of alternately applied a-c and d-c voltages indicated that the ratio of the d-c to a-c breakdown voltage for this stator winding was between 2.12 and 2.3.

9. The points of insulation failure in the breakdown test took place in regions of maximum slot discharge damage.

References

1. DISPERSIONS AND ABSORPTIONS IN DIELECTRICS, PART II, DIRECT CURRENT CHARACTERISTICS, **K. S. Cole, R. H. Cole.** *Journal of Chemical Physics*, New York, N. Y., vol. 10, Feb. 1942, pp. 98–105.

2. TEMPERATURE-RESISTANCE CHARACTERISTICS OF ELECTRICAL INSULATION. *Technical Publication 21-T-4*, James G. Biddle Company, Philadelphia, Pa., 1953.

3. SLOT DISCHARGE DETECTION BETWEEN COIL SURFACES AND CORE OF HIGH-VOLTAGE STATOR WINDING, **John S. Johnson.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1993–97.

4. DIAGNOSES OF A-C GENERATOR INSULATION CONDITION BY NONDESTRUCTIVE TESTS, **A. W. W. Cameron.** *AIEE Transactions*, vol. 71, pt. III, Jan. 1952, pp. 263–74.

5. TESTING ELECTRICAL INSULATION OF ROTATING MACHINERY WITH HIGH-VOLTAGE DIRECT CURRENT, **G. Leslie Hill.** *AIEE Transactions*, vol. 72, pt. III, Apr. 1953, pp. 159–74.

6. TECHNIQUES AND EXAMPLES OF HIGH-VOLTAGE D-C TESTING OF ROTATING-MACHINE WINDINGS, **C. L. Sidway, B. R. Loxley.** *Ibid.*, Dec., pp. 1121–29.

7. THE CHARACTERISTICS OF INSULATION RESISTANCE, **S. Evershed.** *Journal*, Institution of Electrical Engineers, London, England, vol. 52, no. 224, 1913.

◆

Discussion

A. W. W. Cameron (Hydro-Electric Power Commission of Ontario, Toronto, Ont., Canada): The authors have made notable contributions to a subject which is deservedly receiving attention, because it offers substantial benefits in economy of maintenance and reliability of operation of high-voltage rotating machines.

Regarding conclusion 5, I must take issue. I would regard the resistance-versus-voltage curves recorded as good evidence of breakdown prediction. It is all too easy to be wise after the event in this subject, so it would be well to trace the authors' steps carefully.

In Fig. 5, the first test in 10-minute steps, which could not give high accuracy, shows a warning at 20-kv clearly enough to persuade the authors to desist. In view of subsequent breakdown at 24.5 kv, I think that, if this is not a nondestructive prediction, it is a commendable first attempt at it. Trying not to use hindsight, I think that with that curve alone I would guess breakdown at roughly 26 kv.

The next curve, 10-minute steps with negative polarity, is complicated, as the authors say, by reduction in percentage increments, but repeats the ominous downward trend. At 27.5 kv, runaway commences, but complete breakdown at least is averted.

The 30-minute step curve repeats the warning droop. It is unfortunate that there is not a 22- or 23-kv step; no doubt, previous withstanding of 27.5 kv had suggested that 25 kv could now be withstood. I do not quite see why this curve is extrapolated horizontally, observing that breakdown at 24.5 kv occurred during increase of voltage when leakage current could not be measured.

Fig. 7 shows curves for phases *A* and *C* up to 25 kv, which unmistakably announce their impending breakdown with the contrast provided by phase *B*'s 50-kv strength. Considering the droop on *A* and *C*, the 5-kv steps are rather large, and smaller steps over 20 kv might have given a more accurate idea of the actual 26-kv breakdowns. It should be noted that the decrease of resistance between 15 and 20 kv for *A* and *C* is actually less than for *B*. It is the negative second derivative, the droop, that supplies the warning. My experience suggests that a few more points in a case like this might have enabled us to stop at perhaps 23 kv and make a prediction of breakdown at 25 to 27 kv, whereas the available test points suggest 28- to 30-kv breakdowns. Again it may be said that prediction has been perfectly clear, although accuracy might have been improved. The slight downward slope for phase *B* makes me think that its 50-kv breakdown might also have been predicted.

Fig. 9 gives two examples of a type of behavior which I have experienced three or four times in the field and frequently in the laboratory. I did not know of it when I wrote a paper[1] on this subject 3 years ago, but I have mentioned it in discussions since then. A marked and sustained upward "hook" of the resistance-versus-voltage curve has, in my experience, invariably foretold breakdown at a somewhat greater voltage; when a nondestructive test is desired, it is not possible to estimate how much greater.

Thus, in the case of a machine required for service, a curve such as shown for phase *A* would have made me stop the test at 32 kv, and assume d-c strength of the ordre of 33 kv. Considering the narrow difference in physical damage between faults of 33- and 40-kv d-c strength, I submit that such a conservative prediction is good enough for most practical purposes.

I fear I must also disagree with the statement that the decrease of resistance from 37 to 39 kv does not predict failure: to my experience, it does.

This phase-*A* curve is interesting in that, extrapolated from 30 kv, it does trend towards the actual 40-kv breakdown: I have several times noticed that when I have a fair prediction, and am afterwards confused by errative behavior, I was right the first time. Machine condition during these tests, at ambient temperature and slightly humidified, is such as I would consider favorable to prediction.

Regarding the discharge detector, I find that the use of such devices is confused by the sharp current pulses which are associated with the dielectric absorption and dielectric discharge currents, varying in repetition rate with the variation of these currents.

Conclusion 6, regarding the time at which breakdown occurs under a given voltage seems unclear as to how much this depends on deterioration, and how much on the variation of potential gradient in the thickness of the insulation with the progress of absorption. I suggest this as an interesting point for research in connection with direct voltage endurance characteristics.

I hope the authors will become ardent predictionists. I suggest that the first objective is not great precision of the breakdown voltage which we forecast. When a breakdown voltage is forecasted, it is left to our experience and judgment to decide whether or not it represents a physical condition acceptable for continued service. The first objective is to ensure a nondestructive test.

If, as suggested by another discusser, the word "prediction" conveys an unnecessarily high degree of precision, and the word "warning" is more acceptable, I for one will cheerfully accept its use.

REFERENCE

1. DIAGNOSES OF A-C GENERATOR INSULATION CONDITION BY NONDESTRUCTIVE TESTS, **A. W. W. Cameron.** *AIEE Transactions*, vol. 72, pt. III, Jan. 1952, pp. 263–74.

Graham Lee Moses (Westinghouse Electric Corporation, East Pittsburgh, Pa.): This is an excellent analysis of a tremendous mass of test data pointing out some very significant facts. Where direct voltage tests and current leakage data are used to study insulation conditions, it is very important that standardized test methods be employed.

The authors make a point of showing the effects of many variables in testing technique and conditions, which are under the control of the test operator. The curve in Fig. 4 is an excellent example, as the significance of the upper curve could very easily be misinterpreted. This erroneous indication was the result of insufficient discharge time and illustrates the ease with which misinformation can be obtained.

The data indicate that a considerable period of constant voltage is required to eliminate the significant effects of dielectric absorption in the accumulation of d-c leakage data. It is apparent that in most cases when leakage-voltage characteristics are determined, steady-state leakages alone are not measured at each test point. The observed current is usually some combination of steady-state leakage and absorption current. Depending on exactly how the tests are run, the length of time at each test voltage, and other factors, differently shaped curves are quite possible. This raises the question of exactly how such leakage-voltage characteristics should be properly tested. It appears impractical to wait at each test point until the absorptive currents have completely died out. Therefore it seems logical that some reasonable procedure be followed which involves preselected increments of voltage, and fixed time of application at each step, in order to obtain consistent and comparable results. In reporting the results of any given test, the steps of voltage used and the time at each step should always be reported.

Examination of these test results shows that a precise prediction of breakdown voltage was not possible from the insulation resistance versus voltage curves on the complete stator winding or individual phases. There may have been a forewarning in one case, but even this is controversial. This paper re-emphasizes the fact that forewarning and prediction of failure level is often not possible.

---

ol. 74
. 1
955

A New Line of Direct-Acting Recording Instruments

L. J. LUNAS
MEMBER AIEE

J. C. NYCZ
ASSOCIATE MEMBER AIEE

A NEW line of direct-acting recording instruments has been developed to meet the increasing demand for a compact design featuring ease of maintenance, improved operating reliability and conservation of panel space.

Prior to World War II, such instruments were usually projection-mounted with the case removable for routine servicing such as filling the inkwell, renewing the chart, and winding the clock. Consequently, a recorder with 4-inch scale occupied about 6 inches of panel width. The change in switchboard design practice to flush-type mounting of instruments made it necessary to increase the case width to 8 or 10 inches to obtain accessibility for servicing operations, or to retain the 6-inch width at a sacrifice in accessibility.

In recent years the trend toward centralization of electrical and mechanical instrumentation has made it necessary to mount large numbers of instruments within a small area for ease of operation and observation.[1] Thus, the need for minimum space without sacrificing accessibility has steadily increased. To retain the same case width for flush mounting as for projection mounting and yet have convenient accessibility required a new concept in design arrangement. This was accomplished by making all operations and adjustments conveniently accessible from the front. The important operations and adjustments, which have been made fully accessible, are as follows:

1. Renewing the chart paper.
2. Inspection of elapsed record on the reroll spool with automatic rerolling.
3. Filling the inkwell.
4. Adjusting zero.
5. Winding the spring clock.
6. Changing the chart-speed gears.
7. Renewing the illumination lamps.
8. Removing the inkwell for cleaning.
9. Adjusting the movement stops.
10. Adjusting the damping of the mechanism.
11. Changing the indicating-scale plate.
12. Change of calibration after installation for different scales or transformers.
13. Removal of pen for occasional cleaning or for adjustment of pen-tip pressure.

The general appearance of the new instrument is shown in Fig. 1. The instrument proper may be divided into the following components:

1. The case-and-cover assembly.
2. The chart-driving mechanism.
3. The frame assembly, which supports the chart mechanism, scale-plate, and inkwell.
4. The measuring mechanism and base assembly.

Case-and-Cover Assembly

The design of the side-hinged cover has been carefully considered to provide improved visibility of the scale and elapsed record, simple opening and closing with dust-tight gasket seal, and pleasing modern appearance. The case assemblies of the flush- and projection-mounting types differ only in the absence of the flush-mounting flange from projection-mounted cases. Thus, instrument accessibility is equally simple for both types, and instruments can be converted in the field by interchanging case assemblies.

Chart-Driving Mechanism

The chart-driving mechanism is a unit assembly mounted at the lower front edge on bearing supports in the recorder frame assembly. The chart mechanism conveniently tilts forward for accessibility in renewing charts, filling the inkwell, and similar operations. The unit is securely latched in the operating position by a toggle-design lever. Operation of this lever lowers the chart mechanism from the writing pen, thus eliminating the conventional pen-lifter assembly. The indicating-scale plate is mounted on the frame assembly, making it possible to mount the target or indicating pointer of the pen up in front of the scale for good visibility. The chart-driving mechanism incorporates the feed-roll chamber, driving-drum, reroll-spool, reroll-drive unit, and chart-timing mechanism in a single unit, which can be conveniently lifted out of the instrument for maintenance or servicing without the use of tools. The mechanism has been designed to provide chart speeds of $3/4$, $1^{1}/_{2}$, 3, 6, and 12 inches per hour. A 2-speed chart-timing mechanism adds the corresponding speeds in inches per minute.

Inking System

The inkwell is supported in a lever-operated cradle in the frame assembly. The inking system is unique in that the pen is not disturbed when the inkwell is lowered for refilling, accomplished by pulling the operating lever forward. The inkwell is then conveniently accessible for filling without removing it from the

---

Paper 55-484, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer General Meeting, Swampscott, Mass., June 27–July 1, 1955. Manuscript submitted March 24, 1955; made available for printing June 1, 1955.

L. J. LUNAS and J. C. NYCZ are with the Westinghouse Electric Corporation, Newark, N. J.

Acknowledgment is due D. A. Young, of the Westinghouse Electric Corporation, by whom many of the unique principles and features incorporated in the design of the recording instruments were conceived and promoted.

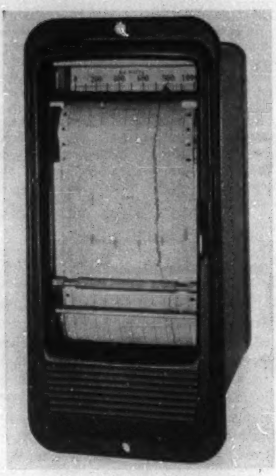


Fig. 1. Front view of new direct-acting recording instrument

instrument as shown in Fig. 2 but can be removed for occasional cleaning by lifting it from the inkwell carrier. The indicating-scale plate is mounted on the inkwell-link system so that it moves up out of the way when the inkwell is lowered.

The writing pen is mounted on knife-edge bearings supported near the front of the instrument on the pen-arm, which is mounted on the measuring element shaft. The pen is easily removable for cleaning or balance adjustment for pen-tip pressure by lifting it out of the pen-arm bearings.

Operating Characteristics

Operating torque of direct-acting recording instruments has heretofore been limited by permissible power consumption, damping problems, and thermal-overload capacity. The availability of high-coercive magnets and recent developments in high-efficiency iron-core magnetic circuits for a-c mechanisms have greatly reduced the effect of these limitations. It was therefore decided to investigate the feasibility and the effect of higher operating torque and to determine the optimum operating torque. In the case of a-c instruments, it becomes practical to double or even triple conventional operating torques, which are now in the order of 100 to 150 centimeter-grams (cmg) for the spring measured over a deflection of 360 degrees. However, it was found that the permissible operating torque for a-c instruments on 60-cycle service was limited by the double-frequency response of the element at higher operating torques. At high chart speeds in the order of inches per minute, this exhibited itself as a sine-wave record of perceptible magnitude, even at torques in the order of 175 cmg for 360 degrees. At lower chart speeds the record line in effect was widened by the oscillation of the moving element and writing system. The problem therefore resolved itself into one of determining optimum torque from the standpoint of pen-tip friction, continuity of inking on quick swings, responsiveness of the element, and tendency of the pen to be pulled toward center from the ends of the pen-arc at high chart speeds.

Tests of pen-to-paper friction made with 20 milligrams pen-tip pressure under vibration-free conditions gave the results shown in Fig. 3. In normal service, particularly on alternating current, the vibration either from a-c excitation of the element or from external sources is normally more than sufficient to eliminate the lower values of friction error. However, the tests indicated the desirability of having operating torque exceed 100 cmg for 360 degrees.

Continuity of inking on quick swings under drastic load changes is necessary for good records. The pen system was designed to develop increased pen-tip pressure on quick swings. This was accomplished chiefly through the arrangement of the pen-bearing support in relation to the center of gravity of the pen. Thus, by placing the knife-edge bearing below the center of gravity of the pen, the writing tip tends to press down or increase writing pressure on sudden swings.

For sudden excursions up to half-scale deflection, the static pen-tip pressure required is substantially zero. For full-scale angle deflections, the required static pen-tip pressure is about 20 milligrams, which is about half the pressure required for conventional pen designs.

The relation between pen-tip pressure and per cent of record skipped or uninked for various operating torques is shown on Fig. 4. This indicated the desirability of operation with torque below 200 cmg for 360 degrees for good continuity of records.

Two important factors governing responsiveness of a direct-acting recording instrument are the inertia of the moving element and the operating torque. The relation is usually expressed as

$$T=2\pi\sqrt{\frac{K}{s}}=\text{undamped periodic time in seconds}$$

where K = inertia of the moving element in gram-centimeters $(\text{cm})^2$ and, s = torque in dyne-cm per radian.

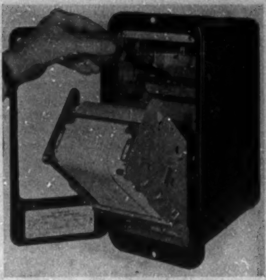


Fig. 2. Inkwell lowered for convenient filling. Pen is not disturbed for this operation

With a moving-element system of given inertia, the natural period varies inversely as the square root of the torque. Actual tests of elements showed that free or natural period varied with torque, as shown in the upper curve of Fig. 5. Under near-critical damping conditions [1-percent (%) overshoot], the full-scale response time varied with torque as shown in the lower curve of Fig. 5. Since the reduction in response time for torque above 150 cmg per 360 degrees diminishes in rate, operation would be well into the region of diminishing returns.

Operation of direct-acting recording instruments at high chart speeds (in the order of 6 inches per minute) gives rise to errors in recording at the scale ends. This is because of the pull-in force on the pen by the chart motion due to the component of force introduced by the arc-shaped locus of pen travel across the chart. Tests run at 25 mg pen-tip pressure over a total scale angle of 50 degrees at 6 inches per minute gave the results shown in Fig. 6. The reduction in pull-in for operating torque over about 150 cmg is not appreciable enough to justify further increasing operating torque as far as this particular factor is concerned.

Thus, investigation of desirable operating torque showed in general that this torque should preferably be not less than 100 cmg per 360 degrees. However, with no limitation in torque imposed by power consumption, overload capacity, or damping, it has been decided that an operating torque in the order of 150 cmg per 360 degrees is as high as is practical. This essentially is the optimum torque from the standpoint of performance factors such as responsiveness, inking, and extraneous friction, and yet it is below the point at which oscillographic response on alternating current causes trouble.

Vol.
Pt. 1
1955

Mechanisms

The three new and different types of measuring mechanisms which have been developed are:

1. Permanent-magnet moving-coil type for d-c, rectifier-type, and transducer measurements.
2. Attraction iron-vane mechanism for alternating currents and voltages.
3. Iron-core electrodynamic mechanism for wattmeters and varmeters.

PERMANENT-MAGNET MOVING-COIL MECHANISM

This mechanism utilizes Alnico-V magnets with a 1-piece laminated pole and air-gap structure. The laminated pole structure is arranged to allow the moving element to be lifted out of the assembly without disturbing any part of the magnetic circuit. This is particularly helpful in making repairs to a moving element.

The mechanism incorporates a unique adjustable magnetic shunt for calibration adjustment. Two iron rods mounted in circular slots in the laminated magnetic circuit as shown in Fig. 7 are adjustable in and out of the stack to shunt, more or less magnetic flux from the operating air-gap portion of the magnetic circuit. This adjustment can be made with the instrument in service and it provides a convenient means of varying calibration to suit different lead resistance for millivoltmeters used as shunt-operated ammeters or to make other adjustments in calibration. A total range of adjustment of about 40% is provided. Instruments are normally shipped with the adjuster near mid-position, which allows ample range for field adjustment. Scale distribution is not perceptibly affected by the shunt adjustment.

The major component of damping is obtained by eddy-current effect in the short-circuited bobbin frame. However, this mechanism, in common with the a-c mechanisms, employs an adjustable permanent magnet and damping disk. This provides an adjustable source of auxiliary damping to suit different circuit conditions. The Alnico-V damping magnets are mounted on a frame plate, which is adjustable in position radially with respect to the disk. The magnet assembly mount is spring-loaded, which simplifies damping adjustment to a screw driver operation as shown in Fig. 8.

The use of Alnico-V magnets and the magnetic circuit design develops operating torque of 150 cmg for 360 degrees with full-scale (50-degree deflection for 4-inch scale) sensitivity of 0.0025 watt. This permits ammeter operation from 50-millivolt shunts and voltmeters of 1,000 ohm-per-volt sensitivity. Ammeters have sufficient zero-temperature coefficient series resistance to reduce temperature influence to be well within the American Standards Association Specification.[3]

For higher sensitivity applications the power consumption can be reduced to 0.001 watt by eliminating the adjustable magnetic shunts and reducing operating torque to 100 cmg for 360 degrees. Such instruments are useful in applications in which the amount of power that can be drawn from the measured source is limited, as in the case of communication circuits.

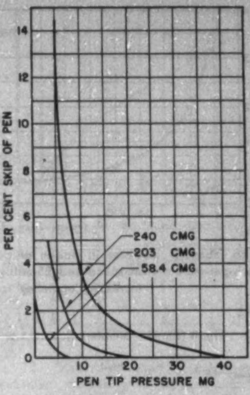


Fig. 4. Relation between pen-tip pressure and per cent of record skipped for full-scale deflection at various operating torques

ATTRACTION-IRON MECHANISM

The attraction-iron principle of operation was selected for the mechanism used for a-c voltmeters and ammeters because of its inherent simplicity, high magnetic efficiency, and high overload capacity. The arrangement of the magnetic circuit is illustrated schematically in Fig. 9. The proportions of the magnetic circuit were determined by the following factors:

1. The geometry of the available space, which governs the proportions of iron and copper based upon the required angular travel of the moving iron and thermal capacity of the winding.
2. Desired scale distribution.
3. Permissible power consumption.
4. Magnetic saturation characteristics. The magnetic circuit was designed to operate well below saturation to minimize wave-form errors.
5. Dynamic performance which is controlled by inertia and torque.

The magnetic efficiency of a moving-iron mechanism may be expressed as follows[2]

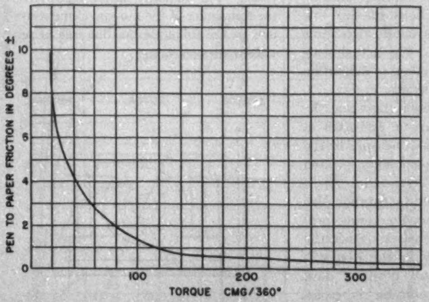


Fig. 3. Effect of pen-to-paper friction under vibration-free conditions with 20-milligram pen-tip pressure and various operating torques, 4.75-inch pen-arm radius

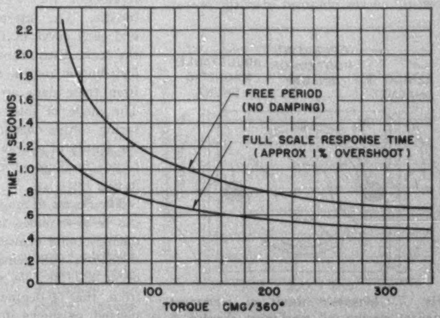


Fig. 5. Variation in undamped natural-period and responsiveness with increasing operating torque

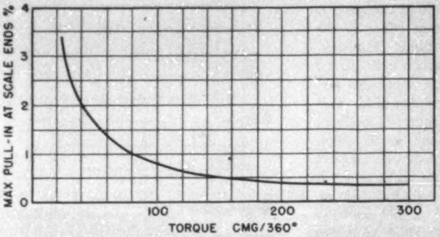


Fig. 6 (above). Decrease in pull-in of pen from scale-arc ends at 6-inch-per-minute chart speed and 25-milligram pen-tip pressure with increasing operating torque

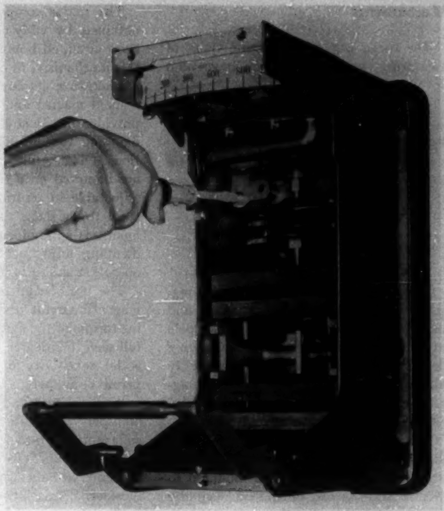


Fig. 8 (right). Simple adjustment of magnetic damping

$$\mu=\frac{L_m-L_0}{L_m}$$

where L_m = inductance at full scale, and
L_0 = inductance at zero

The curve in Fig. 10 shows the inductance characteristics of this mechanism from which the efficiency is calculated to be 35%. Heretofore, magnetic efficiencies have been in the order of 10 to 25 per cent. It can be generally stated that the higher the magnetic efficiency, the lower will be the ampere-turns necessary to develop a given torque. Consequently, higher efficiency permits use of larger size wire, which results in higher thermal overload capacity.

The analysis of performance of any mechanism is facilitated by a study of its torque-producing ability characteristics. These data are obtained by measuring the torque developed for a given number of ampere-turns over the entire range of angular travel of the element. In the case of permanent magnet, moving-coil instruments, and electrodynamic wattmeters with uniform scales, the characteristic curves obtained over the operating ranges appear as families of horizontal parallel lines. However, torque-producing ability curves for this moving-iron mechanism, Fig. 11, are generally representative of instruments having inherently nonlinear scales.

Knowing the desired full-scale torque and angular travel of the moving element, the torque-producing ability curves can be used to select an operating region which will afford the most consistent scale distribution and operating-torque condition. For this mechanism the operating region selected was between the 12-degree and 60-degree positions of the rotor. The straight line drawn from A to B, Fig. 11 represents the torque spring characteristic. The intersections of this line and each curve of the family determine the scale distribution. The actual scale distribution and the method of derivation from these data are shown in Fig. 12. The angle of intersection between the spring torque characteristic and each of the family of curves is a measure of the torque gradient at those points, plotted in Fig. 13.

The degree of magnetic saturation also can be determined from the torque-producing ability characteristics. Since the maximum flux density in the iron will occur at full-scale rotor position, data from the 60-degree position is cross-plotted, as shown in Fig. 14, to indicate the relative degree of saturation. The continuation of the flat region of this curve beyond the full-scale point shows that the iron is operated well below saturation.

The square-law torque characteristic of this mechanism results in extremely high torques during overload conditions. This would normally result in violent deflection of the pen and severe mechanical stresses in the moving element. These effects are eliminated by mounting the rotor iron so that it is free to rotate on the shaft. The rotor is biased against a stop on the moving element by an auxiliary spring. Any sudden overload permits the high torque developed in the rotor to deflect it quickly to the full-scale point. The auxiliary spring then moves the remainder of the moving element system to the full-scale position free of any violent torques.

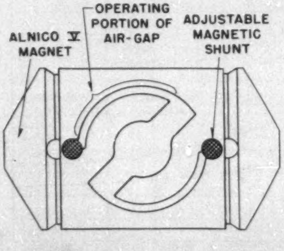


Fig. 7. Magnetic circuit of permanent-magnet moving-coil mechanism showing 1-piece lamination and adjustable calibration shunts

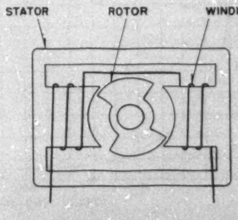


Fig. 9. Magnetic circuit of attraction-iron mechanism for alternating voltage and current measurements

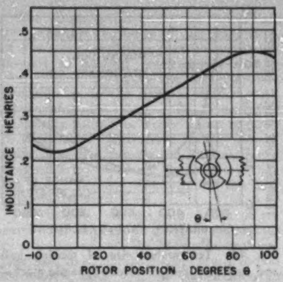


Fig. 10. Variation of inductance of moving-iron mechanism with rotor position

Power consumption is as follows:

The 0–5 ampere ammeter at 5 amperes, 60 cycles: 0.4 watt, 3.4 vars, 3.5 volt-amperes, and 10% power factor.

The 0–150 volt voltmeter at 120 volts, 60 cycles: 14.2 watts, 1.6 vars, 14.4 volt-amperes and 99%+ power factor.

Temperature and frequency influences are well within the specifications.[3]

IRON-CORE ELECTRODYNAMIC MECHANISM

The highly efficient, close magnetically coupled iron-core electrodynamic mechanism has been used in indicating instruments in this country for the past 15 years. The technique of compensating for moving-coil inductance and iron losses, which normally cause power factor and frequency errors, has been perfected and generally used. The high magnetic efficiency producing high operating torques at relatively low power consumption makes this type of mechanism particularly suitable for the high torque requirements of direct-acting recording instruments.

The mechanism as it has been developed for this recording instrument has the magnetic circuit built up of 1-piece lamination, thus insuring an inherent, consistently uniform scale of linear distribution.

The magnetic circuit structure is built up as illustrated in Fig. 15 with one half of the laminations of each element reversed with respect to the other half so that the inductance of the moving coil remains constant over the operating range. This eliminates the solenoid effect which would otherwise cause deflection on voltage alone and also excessive voltage influence. Another unique feature of the voltage element is the splitting of the moving coil of each element into two separate coils, which permits the use of the astatic connection, in which the current is made to flow in the active sides of both coils in the same direction. This cancels the effect of uniform stray fields upon the mechanism.

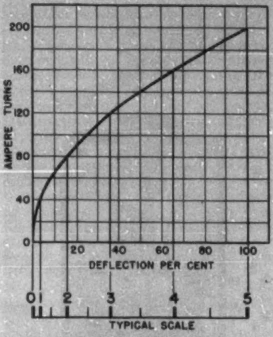


Fig. 12. Scale distribution of moving-iron mechanism derived from torque-producing ability curves of Fig. 11

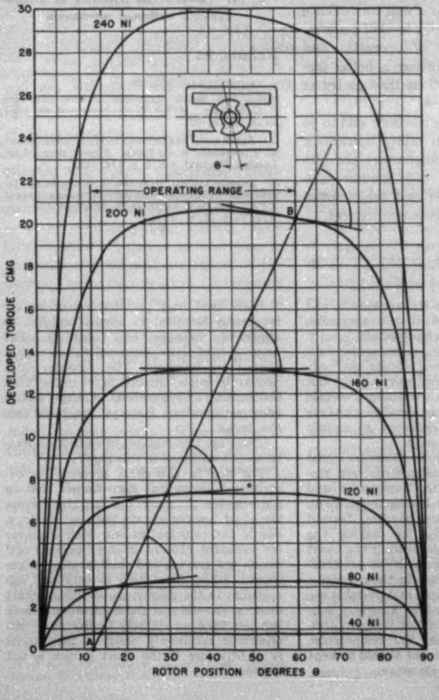


Fig. 11 (left). Torque - producing ability curves of attraction-iron mechanism for various ampere-turns. Methods of deriving scale distribution, operating torque gradient and saturation characteristics indicated

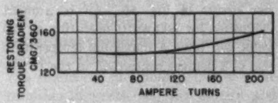


Fig. 13. Cross-plot from Fig. 11 deriving restoring torque gradient

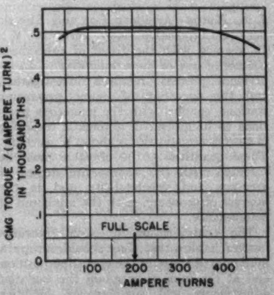


Fig. 14. Cross-plot from Fig. 11 deriving relative saturation characteristics

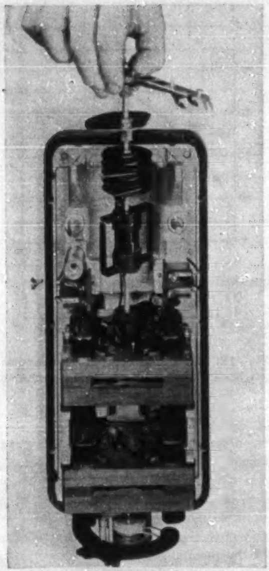


Fig. 15. Electrodynamic mechanism showing 1-piece laminations and structural arrangement

The superior performance of this new mechanism is shown by the temperature rise curves of Fig. 16. An instrument calibrated for 500 watts operating at 120 volts will have a maximum temperature rise of 8 degrees centigrade in the potential coil and 6.5 degrees centigrade in the current coil. The low temperature rise of the coils makes it possible to change calibration over a wide range of scale conditions without the necessity of changing the coils or otherwise rebuilding the mechanism.

The compensation for varying power factor is accomplished in the conventional manner by connecting a suitable capacitor across a portion of the series resistance of the voltage circuit. The performance of this instrument under various conditions of power factor is shown in Fig. 17.

Power consumption of the 5-ampere 120-volt 2-element wattmeter is as follows, per element:

Current circuit at 5 amperes, 60 cycles; 6.0 volt-amperes, 1.0 watts, 5.9 vars, 17% power factor.

Voltage circuit at 120 volts, 60 cycles; 5.8 volt-amperes, 5.8 watts, 0.0 vars, 100% power factor.

Conclusions

The development of this new line of recording instruments represents a marked advanced in performance and convenience of operation. This high level of performance and simplicity of operation and maintenance should assure a better picture of system operations through better records.

Control-room space has been said to be the costliest space in utility operations. This new recording instrument provides better performance and greater convenience of operation and maintenance with a minimum of space requirement.

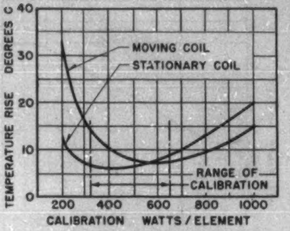


Fig. 16. Temperature rise of moving coil and stationary coil of electrodynamic mechanism at rated voltage and for currents determined by the various watt calibrations

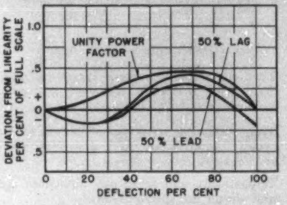


Fig. 17. Power factor influence of electrodynamic wattmeter mechanism

References

1. Centralized Instrumentation and Controls for Steam Electric Power Stations, B. C. Mallory. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 591-97.

2. Electrical Measuring Instruments (book), C. V. Drysdale, A. C. Jolley. Ernest Benn, Ltd., London, England, vol. I, p. 276, 1924.

3. Direct-Acting Electrical Recording Instruments (Switchboard and Portable Types). *ASA C39.2-1953*, American Standards Association, New York, N. Y., 1953.

◆

Discussion

J. A. Morris (American Gas and Electric Service Corporation, New York, N. Y.): In keeping with the trend of centralized control of one or more functions of an individual process, research investigation, and the operating of power generating stations, coupled with the numerous functions to be measured, automatic recording of electrical quantities is almost a necessity. These quantities to be of value must be recorded with the appropriate degree of accuracy and dependability and in the interest of economy, with the minimum of maintenance.

The centralization of such recorders further implies that they should occupy a minimum of space in the operating position and at the same time retain the essential adjustments and maintenance facilities easily available.

A certain degree of freedom which permits changing the operating characteristics, such as scale calibration, chart speeds, and damping, is desirable if the maximum flexibility of usage is to be obtained.

In the portable form, recorders for voltages, current, and power should not only have the desirable features noted, but should also be able to withstand a certain amount of abuse without significant change in operating characteristics or to sustain permanent physical damage.

The designer of such recorders must recognize that the instruments are frequently, if not usually, transported and installed by personnel having a minimum of training in instrument handling beyond the essential knowledge of making the external connections and setting the chart time. Therefore, ruggedness of construction is an essential factor in their design.

Recorders both for fixed and portable use are frequently installed in locations subject to uncontrolled ambient temperatures. This fact requires that the temperature coefficient of the instrument should be kept to a minimum over a wide range.

Both fixed and portable recorders are frequently installed in circuits subject to abnormal voltages and currents. The occurrence of these abnormal values cannot always be foreseen. The recorder connected to such circuits may suffer both mechanical and electrical damage unless generous overload features are built into the device.

There are two features of recorder operation not mentioned in this paper.

The first is the chart drive. The experienced instrument man knows that in some devices the power available for driving the chart has not always been satisfactory, due to two principal causes: lack of power for extended (15- to 30-day) periods and poorly cut gearing and shaft bearing alignment. While these two defects are less apparent in the electric-motor-driven chart devices, the story of the hand-wound clock has not been an entirely happy one. An otherwise excellently designed instrument is of little use if the chart device is not dependable.

The second point is the ink employed. The desirable attributes of this item, some of which are conflicting, may be mentioned in part as follows: quick drying on the chart; slow drying in the ink reservoir; noncorrosion to the parts in contact with it; nonfading; sharp record contrast; and easily soluble in either water, alcohol, or petroleum derivative. While all of these desirable features are difficult if not impossible to obtain, it is necessary that the matter receive attention and the best compromise worked out.

We believe that the new design of recorders, as outlined in the paper, provides practically all of the desirable features discussed to a remarkable degree and should be considered a valuable contribution to the instrumentation field.

J. C. Nycz and L. J. Lunas: The history of recording instruments has shown that Mr. Morris's comments are very well taken, particularly in regard to reliability of hand-wound clocks and the problems of recording ink systems. The reliability, accurate time keeping, and low maintenance and servicing of synchronous-motor-driven clocks makes them desirable wherever a continuous, reliable source of power is available. However, if outages occur, it is necessary to use hand-wound spring clocks, which, as stated by Mr. Morris, require considerably more maintenance and servicing. In the design of this recording instrument, a hand-wound clock has been provided which has a 50-percent increase in output torque over previous designs, which should greatly increase the reliability of the chart drive under severe service conditions. Furthermore, the escapement is temperature-compensated to provide accurate time keeping over a wide range of temperature conditions, such as would be encountered in portable service and in unattended substations.

The problems of recording mediums has been under investigation for many years in an attempt to find a substitute for liquid ink. However, with all the problems accompanying the use of liquid ink, no answer has been found to substitute for ink under the widely varying load conditions incurred in the use of electric recording instruments. Improvements have been made in inks, but as stated by Mr. Morris, the requirements are conflicting in the sense that the ink must be slow drying in the inkwell and yet quick drying on the paper chart, so a compromise is always necessary. In the design of this instrument, these shortcomings have been recognized, and the inking system has been designed to minimize the evaporation of ink from the well by designing the inkwell cover with a baffle around the pen opening. From an operating point of view, the chief feature that has been incorporated is to design the inking system for greater ease of filling the inkwell, making this periodic maintenance operation simple and convenient.

Magnetic Amplifier Sequential Circuits of the Asymmetrical Type

ALAN S. FITZGERALD
MEMBER AIEE

THE intent of this paper is to disclose as briefly as possible some new ways of interconnecting magnetic amplifier flip-flops in the hope that these can be adapted for useful purposes. These circuit arrangements may be considered as analogues of manually or mechanically actuated contact-making devices. The basic switch function is an operation in one sense for ON and an action in the opposite sense for OFF as in the common toggle switch. Another familiar type of switch, not quite so frequently encountered, is called the "canopy" or "chain-pull." When this is operated successively, in one sense only, it gives alternately ON and OFF action. Relays that do this are well known and are sometimes called impulse relays. When relays of this type are repeatedly energized, they will alternately open and close their contacts. Devices capable of this kind of response form the basis of binary systems applicable in the digital computer field.

The author has attempted to devise systems, analogous to the latter type of switch action, utilizing magnetic amplifier flip-flops. It is apparent that the circuit principles shown are equally applicable to other instrumentalities such as relays, electron tubes, transistors, etc.

The expression asymmetrical is used to distinguish the systems described from those circuits which require duplicate balanced loads such as the Eccles-Jordan tube circuit[1] or the magnetic amplifier arrangements of Thompson[2] and Wood.[3] The application of the asymmetrical type of magnetic amplifier flip-flop to digital computers has been described by Miles.[4] The circuitry used by Miles is economical of material but would appear to be somewhat critical as regards the exact characteristics of the count signal pulse.

The first of the two circuits here described aims at a somewhat more rugged and positive type of action, an exact analogue of the type of relay referred to in the foregoing. That is to say, each time the system is energized, the resulting response is to be opposite to that which occurred previously, regardless of the duration of either the period of interruption or the period of energization.

This result has been accomplished by making use of a characteristic of the magnetic amplifier flip-flop which, though it has been recognized, has not, to the author's knowledge previously been exploited.

Interrupted Supply Control

The problem to be solved is readily apparent if the basic characteristics of the magnetic amplifier flip-flop[6] are considered. These are more fully described in reference 6. As there explained, when enough feedback is incorporated in a magnetic amplifier, the system has bistable characteristics. By providing the proper bias excitation it can be made to remain indefinitely in one or the other of two conditions of stability, one of these resulting in a relatively low current and the other giving a current of substantially higher value. The magnetic amplifier may then be caused to transfer from either of these stable states to the other one by the momentary application of a magnetomotive force (mmf) applied by a control winding. Normally, the current in this control winding is of reversible polarity. Herein the transfer to the high-current state will be called "flip," and reduction to the low-current condition "flop."

Note that it takes controlling effects of dissimilar type to cause flip and flop action respectively. So a system like the common toggle switch results, with action in one sense for flip and action in the opposite sense for flop. What is desired is the chain-pull type of response. A series of control signals of unchanging polarity should cause alternately opposite responses.

This result can be achieved by making use of an interesting characteristic in-

---

Paper 55-620, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Pacific General Meeting, Butte, Mont., August 15-17, 1955. Manuscript submitted April 26, 1955; made available for printing June 23, 1955.

ALAN S. FITZGERALD is at 333 Corte Madera Avenue, Mill Valley, Calif.

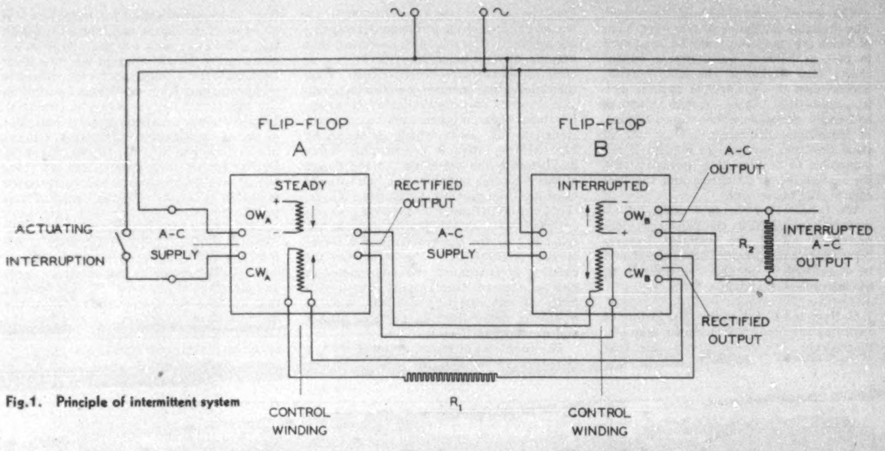


Fig. 1. Principle of intermittent system

herent in the magnetic amplifier flip-flop. As has been explained in the foregoing, the system can be caused to transfer from the flip condition to the flop value, and vice versa, by the application of a control mmf of the right polarity. Such a control signal must of course be of an adequate order of magnitude to accomplish this result. This, of course, refers to the condition in which the flip-flop is normally energized with alternating current. If a control signal of insufficient amplitude is applied, no action will result under this condition.

But if such a reduced control signal is applied with the flip-flop de-energized, and this signal is maintained, and the alternating voltage is applied to the flip-flop, it is found that the polarity of this weak control signal is capable of predetermining which of the two conditions of stability, flip or flop, results when the alternating voltage is applied. As long as the flip-flop remains energized, however, any further change in polarity of the low-level control signal is without effect.

How this can be done by interconnecting two flip-flops is shown in Fig. 1. As may be seen in Fig. 9 of reference 6, the flip-flop per se is likely to include a-c windings, feedback windings, bias windings, and rectifiers, regardless of the method by means of which the flip-flop is to be actuated. To simplify the explanation of how the system works, those portions of the circuit structure which are conventional elements in a flip-flop are omitted in the block diagram. The exact magnetic amplifier structure of the flip-flop is not of primary significance in the present sequential circuit arrangement. Any kind of magnetic amplifier that will flip-flop can be used. Thus, in Fig. 1, it is assumed that the rectangles represent any suitable type of magnetic amplifier including such cores, a-c windings, feedback windings, bias windings, and rectifying elements as may be necessary to provide the proper flip-flop effect. Terminals for receiving the a-c supply, and output terminals, are shown in the rectangle representing each flip-flop. The only windings actually depicted are those for applying the controlling mmf in accordance with the action of the system.

Thus Fig. 1 shows two flip-flops, *A* and *B*. They are neither necessarily nor desirably of similar magnitude and have no symmetrical relation. Flip-flop *A* delivers rectified unidirectional output, the flip value being substantially greater than the flop value. The latter, however, may be displaced from zero to any extent that may be convenient.

In a magnetic amplifier flip-flop circuit there is usually a rectified current with flip and flop values for supplying the feedback winding. There is also a corresponding alternating current likewise having flip and flop values—the same currents before rectification. The usable output current may therefore be withdrawn either as a-c output, or rectified output, or both. Accordingly, in Fig. 1, the right-hand flip-flop *B* is shown with two sets of output terminals, one for a-c output and one for rectified or d-c output.

One other special feature is to be included within the rectangle representing flip-flop *B*, namely, the flop value of the output, both a-c and rectified, is to be treated as a quiescent current in a magnetic amplifier. A suitable compensating arrangement is to be incorporated so that the flop value is within a small percentage of zero. Methods for compensating for quiescent current have already been published.[7]

Now consider the arrangements necessary for each flip-flop to be given a controlling input derived from the output of the other. To control either of the flip-flops, an mmf of reversible polarity has to be furnished. But neither of the flip-flops puts out a reversing output. Flip-flop *A* gives d-c flip and flop currents of differing magnitude and of the same polarity. Flip-flop *B* gives a flop value of zero but it does not furnish a reverse polarity. Therefore, an appropriate structure for deriving reversing mmf from the outputs available must be included. This is easily done by providing duplicate control windings in opposing sense. One carries the flip and flop values of rectified output current from the other flip-flop. The other carries a current of a single predetermined magnitude, giving, in opposite sense, a mmf of approximately median value between the said flip and flop magnitudes. Thus the resultant mmf will reverse polarity as desired when the output current of the other flip-flop changes from one value to the other.

As shown in Fig. 1, the output of *A* is connected directly to the control coil CW_b of *B*. The d-c output of *B* is connected to the control coil CW_a of *A* through a relatively high resistance R_1 so as to limit the current to a low value. To indicate the sense of these controlling mmf, an arrow points upwards to denote a mmf in the direction such as to cause transfer from flop to flip, and vice versa.

Vol.
Pt. 1
195

Thus in the case of B, the connection is made in the inverse sense, i.e., whenever A flips, B should, positively, be caused to transfer from flip to flop. Likewise, when the output current of A changes from the flip value to the lower flop value B should be caused, as a result, to flip. This is arranged by connecting the control coil of B in a sense so as to cause the flop response. Thus, whenever A flips, i.e., when the output current of A increases from the low to the high value, the mmf in the control coil of B preponderates over that of the opposing winding OW_b and causes B to flop. Conversely, whenever A changes from flip to flop, delivering the low output current to the control coil of B, the mmf of the opposing winding OW_b of B is greater than that of the control coil CW_b so that the resultant mmf is in the flip direction. Thus, at all times the action of B responds directly, but in opposite sense, to the action of A.

The control winding CW_a of A, on the other hand, is connected so that the mmf is in the direction such as to cause transfer from the flop condition to the flip condition i.e., to cause the output current of A to change from the low value to the high value. Thus, when B is flipped, the mmf caused by CW_a exceeds that of the opposing winding OW_a and the net result is in the direction to cause A to flip. Similarly, when the output current delivered by B has the flop value—which, actually, is zero—the resultant mmf is that caused by the opposing winding alone, which is in the flop sense.

However, because of the resistance R_1, these mmf are of insufficient magnitude to produce any effect upon A under normal conditions i.e., when A is energized from the a-c source. But if flip-flop A be de-energized, when the power is subsequently reapplied, A will assume a stable condition corresponding to the sense of the resultant control effect derived from B. Flip-flop B will instantly respond in the inverse sense. At this point further action is blocked because of the effect of R_1. It is necessary that the a-c supply to flip-flop A be withdrawn and again restored to produce further response.

In Fig. 1 the supply terminals of flip-flop B are shown permanently connected to the a-c source. Flip-flop A, however, is energized through a switch by means of which the a-c power to A may be interrupted and restored at will.

Suppose, for example, that this switch be closed and that the system is at rest with B in the flip state. Under this condition the sense of the signal applied to A is in the direction so as to cause A to flip. But because of R_1 this has not yet occurred. If, now, the switch be opened and closed, A will flip and the system will come to rest with B in the flop condition. A second operation of the switch will restore the original condition with B flipped.

Thus, successive operation of the switch applying a-c energy to A gives rise to the sequential operation desired. It takes two operations of the switch to complete one flip-flop cycle. Thus the system constitutes a binary in which the number of operations of the switch is the count signal. However, as stated previously, this system is based upon a positive blocking action. It does not require that the actuating signal be a pulse of any specific character and it is entirely unaffected by prolongation of the periods, either of energization or interruption.

There is one item of the structure indicated in Fig. 1 that requires mention. This refers to the energization of the two opposing control windings OW_a and OW_b. Flip-flop B is continuously energized. Accordingly, the conditional control signal emanating from B is at all times present in the control coil of A. True, the flop value is zero but there is no condition in which the controlling effect of B upon A is intended to be withdrawn. Thus, as noted on the diagram, current in OW_a may be maintained at a steady value by any convenient means.

The flip-flop A, on the other hand, is not continuously energized. The output current derived from A is only present in the control coil of B when A is connected to the a-c supply. The mmf due to the opposing control winding of B is not desired in the absence of the control signal from A. It would cause an undesired flip action. Thus the opposing winding OW_b is to be supplied with rectified current of the desired magnitude through means derived from the a-c supply terminals of the flip-flop A, which are only energized when the switch is closed. It is, therefore, indicated in the diagram as "interrupted." Thus the currents in CW_b and in OW_b are simultaneously present or simultaneously absent.

Output from the sequential control circuit may be withdrawn from the flip-flop B for any desired purpose either before or after rectification. In Fig. 1, a load circuit represented by a resistance R_2 is shown connected to receive a-c energy from B. The reason for the requirement that the flip-flop B have a flop value close to zero is now apparent. If a duplicate of the circuit of Fig. 1 be connected so that the supply terminals of the flip-flop A of the second unit are energized from the a-c output of the flip-flop B of the first unit, the second unit will be controlled by the output of the first unit in the same way the first unit is actuated by the switch. The second binary stage will then give a count of four. Other successive stages may be added in accordance with the accustomed practice in the binary digital art.

It will be apparent from the foregoing that rather than there being any symmetrical relation between A and B, since the output of B supplies a-c power to the flip-flop A in a succeeding stage, flip-flop A is not of commensurate power rating. It has only to supply an output power level of the order required by a magnetic amplifier control input.

A practical detail of some interest arises from the problem of a flip-flop supplying a flip-flop. It might well be anticipated that the wide variation in the impedance represented by the flip-flop A, considered as a load circuit to be supplied by flip-flop B, could be a source of difficulty. As noted in regard to Fig. 19 of reference 6, varying the load impedance is one way of actuating a flip-flop. This circumstance does of course require the presence of the load resistor R_2. It is an advantage that vectorial rather than arithmetical relations are involved. In practice excessive power wastage in R_2 does not take place. If the resistance of R_2 be of the same order of magnitude as the impedance of the flip-flop A, the dynamic activity of the flip-flop A, as a load, causes no serious functional distress to the flip-flop B from which it is energized, and no difficulty is encountered on this score.

CIRCUIT DETAILS

In the preceding sections there have been omissions of some essential structural matter. Fig. 2 shows more particular data.

No importance or significance is to be attached to the particular types of magnetic amplifier circuits employed. As stated, the meat of the circuitry lies in the interrelation between the flip-flops, not in the kind of flip-flop. This work was carried out a number of years ago and the types of magnetic amplifier circuits shown in Fig. 2, for various fortuitous reasons, happened to be those most readily accessible to the author at that time. Since then magnetic amplifiers of superior characteristics have become available. Thus Fig. 2 shows one way that works, and, as such, may be studied by those who have special interest in the circuit structure. The system will undoubtedly work better if more up-to-date magnetic amplifier circuitry be incorporated in the flip-flops themselves.

The method of compensating for the

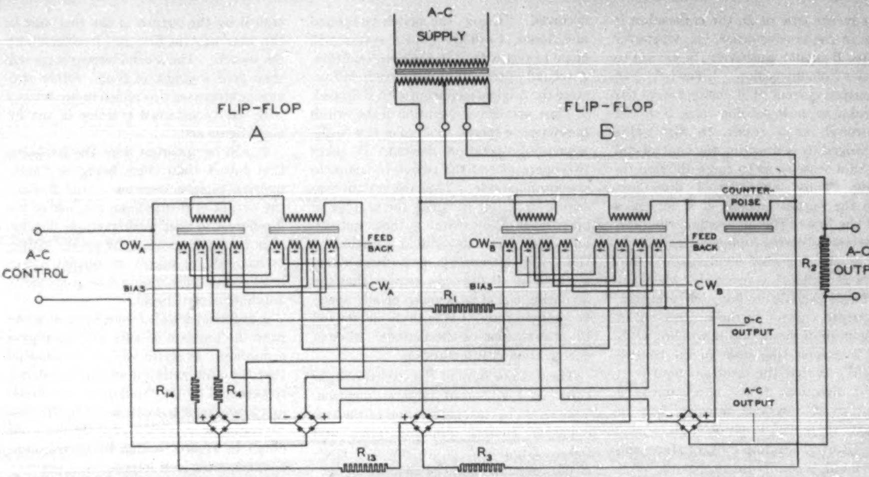


Fig. 2. Circuit details of intermittent system

quiescent, or flop, value of the output current of the flip-flop *B* is shown in Fig. 1(B) of reference 7. This requires a 3-wire a-c source, not unusual in magnetic amplifier circuits. It is believed the principle of operation of Fig. 2 can be followed more easily according to this arrangement than if the less familiar compensating circuit of Fig. 1 (D), reference 7, were shown, which requires only a 2-wire supply and is equally applicable.

In Fig. 2 the control windings CW_a and CW_b, and the opposing windings OW_a and OW_b, are identified as in Fig. 1, as are also the resistances R_1 and R_2. All other circuit elements shown in Fig. 2 pertain to the structure of the flip-flops as such. It is to be noted that the provision of rectified current to the opposing windings OW_a and OW_b requires no additional material because both flip-flops, to retain either their flip or their flop state in the absence of any control signal, require a bias winding. Obviously, since flip-flop *B* is energized with alternating current all the time, the bias excitation must likewise be maintained. On the other hand, since flip-flop *A* is intermittently energized, clearly the bias should not be present when the flip-flop is not receiving a-c energy. Thus, the bias circuit of each flip-flop has the desired energization requirements of the opposing windings on the other. Accordingly, they may be series connected and the turns appropriately proportioned. Note that the excitation of CW_a through R_1 can be derived from the voltage drop across the feedback winding of the flip-flop *B*.

Fig. 3 is a block diagram showing how three binary elements such as the circuit of Fig. 2 would be connected together for applications in the field of digital computing. This would give a count of eight. The methods of adapting such arrangements to count in decades are conventional in the digital computer art and need no further discussion. It is to be understood that in Fig. 3 the count signal consists of intermittent a-c energy applied where indicated.

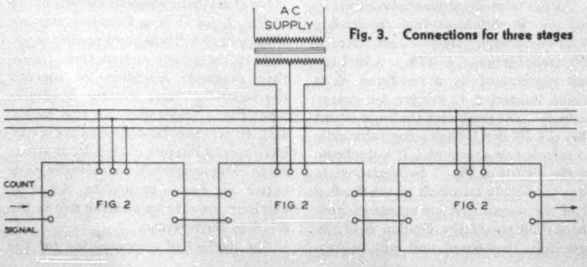


Fig. 3. Connections for three stages

Pulse Signal Control

It has been explained that the objective in Fig. 1 was to get ON-OFF-ON sequential response from actuating electrical effects of a somewhat indefinite nature as to time characteristics; or as distinguished from momentary pulse or trigger signals which necessarily have to be of rather definite duration and amplitude, and uniform within very close limits.

While the previous circuit is of some technical interest in that it is capable of direct response to a count signal characterized by a very slow rate of increase and diminution, it is not difficult to contrive trigger devices which will generate pulse signals from most of the electrical or other effects it may be desired to count.

Accordingly, it would seem of interest to consider the possibility of adapting the cross-connection principle of Fig. 1 so as to obtain a circuit that can be operated by pulse signals instead of by keying the a-c supply to a flip-flop. Such an arrangement is shown in Fig. 4.

Flip-flops *A* and *B* are used as in Fig. 1. These have the same general characteristics respectively. That is to say, the output of *A* is unidirectional and it is not necessary that the flop value be zero.

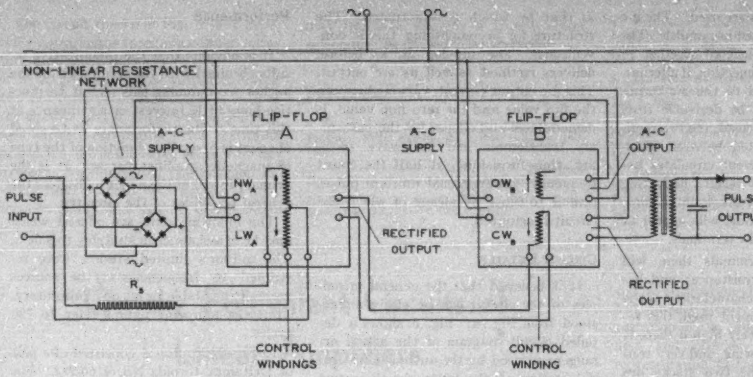

~
~
NON-LINEAR RESISTANCE NETWORK
A-C SUPPLY
FLIP-FLOP A
NW_A
LW_A
PULSE INPUT
R_5
CONTROL WINDINGS
RECTIFIED OUTPUT
A-C SUPPLY
FLIP-FLOP B
OW_B
CW_B
A-C OUTPUT
PULSE OUTPUT
RECTIFIED OUTPUT
CONTROL WINDING

Fig. 4. Principle of pulse system

Flip-flop B, likewise, delivers both a-c and d-c output, and it is desirable that the flop current have a value very close to zero. The d-c output of A is also connected, as before, in reverse sense, to the control winding CW_b of flip-flop B. The opposing winding OW_b on B is again provided.

However, in Fig. 4, both of the flip-flops are connected solidly to the a-c supply. The control coil CW_b is at all times excited by the output current of A without interruption, at times with the larger flip value, and at other times with the lower flop value. Thus in Fig. 4, the flip-flop B carries out a somewhat more simple function than in Fig. 1. All that is required now is that it respond positively, in inverse sense, to the action of A; that it shall instantly flip when A flops and flop when A flips. It does not have to retain a stable flip condition nor hold the flop state, with the control current withdrawn, as was the case in Fig. 1. But it does have to be a true flip-flop. A conventional magnetic amplifier with a more or less linear response is not indicated for this function. No intermediate condition of stability between the flip and the flop states should be possible.

The opposing winding OW_b is supplied with uninterrupted rectified current of a fixed value so that, when CW_b receives the flop value of the output current of A, the resultant mmf will flip B and, when A delivers the flip value, B will flop.

To actuate the flip-flop A, as before, it is necessary to furnish a controlling mmf of reversible polarity. The input signals are assumed to be unidirectional pulses of unchanging polarity. Therefore, a device is needed which will convert the uniform pulse signals into pulses of alternately reversing sense or polarity.

This is accomplished by providing the flip-flop A with two duplicate control windings, opposed in sense, both of them simultaneously energized by the incoming pulse signal. The signal input circuit is divided so that two parallel paths exist. One of these includes the control coil LW_a and a resistance of fixed value R_5. The other path energizes the second coil NW_a through a resistance which is not of fixed value.

The turns of LW_a and NW_a are so proportioned that, if the value of the fixed resistance is greater than that of the variable one, the signal pulse, dividing and passing through both of the control windings, causes A to flip. Conversely, if the value of the variable resistance exceeds that of the fixed one, the signal impulse flops A.

If the value of the variable resistance could be caused to increase and decrease alternately with each successive input pulse signal, A could be made to flip and flop with alternate incoming signal pulses. It is possible to do just this.

A nonlinear resistance element is used as a variable resistor. It is well known that most types of rectifier exhibit marked nonlinear characteristics. As shown in Fig. 4, two conventional full-wave or

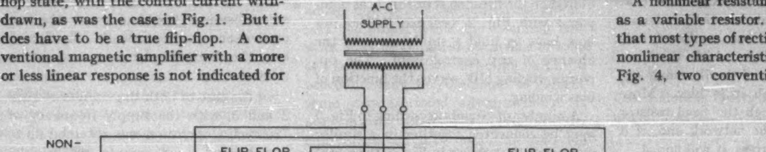

A-C SUPPLY
NON-LINEAR NETWORK
FLIP-FLOP A
FLIP-FLOP B
COUNTER-POISE
R_6
PULSE INPUT
FEED BACK
OW_B
FEED BACK
TRANSFORMER
PULSE OUTPUT
BIAS
NW_A
LW_A
R_8
CW_B
R_{17}
D-C OUTPUT
CHOKE
R_5
R_7
A-C OUTPUT

Fig. 5. Circuit details of pulse system

bridge rectifier stacks are used. The a-c terminals are connected in parallel. The d-c terminals are connected together $-$ to $+$. With this connection, if alternating current is applied to the a-c terminals, no output will be derivable from the d-c terminals. Instead, the two rectifiers are short-circuited between themselves and direct current circulates between them. As a result, neglecting filterable effects, no voltage, either unidirectional or alternating, will appear at these short-circuited d-c terminals.

Between these terminals there will exist nevertheless a resistance; and because of the nonlinear characteristic of the selenium or other element used, this resistance will be relatively high if there is little or no current flowing, and very considerably lower if the two stacks are energized with alternating current so that a substantial direct current flows through the short-circuit path. The path lying between the two pairs of conjoined d-c terminals, is considered to constitute the variable nonlinear resistance.

The value of this variable resistance has to be controlled at will to produce the desired result. This can be readily done by appropriately energizing this nonlinear network with alternating current. The basis of the method of connecting the two rectifier stacks is that this arrangement results in decoupling between the a-c and the d-c circuits, as in a magnetic amplifier. The network, of course, is not intended to act as an amplifier and does not do so. But no electromotive forces emanating from the source from which the nonlinear network is energized appear in the signal input circuit.

Accordingly, the nonlinear network is energized with alternating current derived from flip-flop *B*. Any time that *B* is in the flop condition, there is no a-c energy supplied to the nonlinear network. The resistance which this presents to the portion of the incoming signal passing through it is high. More current flows through the fixed resistor R_s than through the network and, if a signal pulse now arrives, it will flop *A*.

Instantly *B* flips. The network is now energized with alternating current and presents a low resistance value to the next incoming signal pulse. The current through R_s is now less than that flowing through the variable resistance. Accordingly, the next signal pulse flips *A*, whereat *B* is flopped and so on.

This concludes the operation of a single sequential circuit. However, if such a system is to be used as a binary for digital computer purposes, the circuit must emit a signal of the same character as that by which it is actuated. The structure for accomplishing this is conventional. The flip-flop *B*, as before, delivers rectified as well as a-c output. This d-c output current, varying between the flip value and the zero flop value, is demodulated and clipped by the customary transformer and half-wave rectifier, thus furnishing, at half the count frequency, unidirectional uniform pulses similar to those by means of which the circuit is actuated.

CIRCUIT DETAILS

It is believed that the general principles of the circuit can be easily understood from Fig. 4. Fig. 5 shows a detailed circuit diagram of the actual arrangement used by the author in his initial studies of the pulse signal control method.

As in Fig. 2 the counterpoise system is used to obtain the zero compensation for flip-flop *B*. No significant reasons are to be inferred from the use of the customary self-saturating type of magnetic amplifier in one instance and not in others. The author used the most direct approach to securing the desired sequential action as such, rather than making a study of the most effective circuit design for any practical application, or any particular speed of count.

In Fig. 5 the nonlinear network is shown as in Fig. 4. The fixed resistor R_s as well as the windings LW_a, NW_a, CW_b, and OW_b are similarly designated. All other resistors in Fig. 5 serve functions which are merely quantitative. The filter capacitors and choke are conventional. All other windings are normal components of magnetic amplifier flip-flops. One d-c winding less is required on the flip-flop *B* in Fig. 5, as compared with Fig. 2, because *B* does not now have to hold a stable state in the absence of any control mmf. The opposing winding OW_b serves the function of bias winding.

A number of circuits according to Fig. 5 may be connected together in a similar manner to that shown in Fig. 3, the count signal now being a brief unidirectional pulse instead of interrupted alternating current.

Though outdated as regards magnetic amplifier types, this arrangement was fully functional. The actual means employed are authentically disclosed to give a convenient starting point for using this circuit principle in association with the more effective high-speed magnetic amplifier elements now available to which the author did not have access when engaged in this development.

Performance

It is obvious that the quantitative results obtained with the setup used by the author for reducing this circuit to practice have little interest or significance at this time. The performance as to speed of count is basically a function of the type of magnetic amplifier per se. It is the primary intent of the paper to disclose the general principles of the circuitry.

The following data are offered which may be of assistance in studying this field. The author's limited studies were restricted to frequencies of the sources normally available in his own laboratory. Thus the following figures refer to 720 cycles.

In Fig. 2, flip-flop *A* consisted of a pair of Deltamax toroids No. *4168-D2*. For flip-flop *B*, No. *5387-D2* were used. In Fig. 5 both flip-flops were No. *5387-D2*. As to the counterpoise elements, while it is normally assumed that these should comprise magnetic structure having congruency with the magnetic amplifier cores, for reasons not fully apparent to the author, it was found that the flip-flop *B* could be compensated so as to give outputs, at flop, of values sufficiently close to zero by means of counterpoises which were completely resistive.

Both systems operated very positively at relatively low speeds of count. As is to be expected, when the speed of count is increased, the nature of the response necessarily tends to become more and more dependent upon the character of the count signal. For example, in the case of the circuit of Fig. 5, the response is affected by the pulse amplitude and by the pulse duration considered as a percentage of the pulse frequency cycle. When the circuit of Fig. 2 is operating at any substantial speed of count, the successive energizations of the flip-flop *A* have the character of a pulse, and similar relations exist.

In the case of both the circuits of Figs. 2 and 5, with the supply frequency of 720 cycles, operation was obtained up to and beyond 1,500 counts per minute. It would thus appear that, with the types of magnetic amplifier used, the limit of counting speed is reached when there is available for, response and recovery of themagnetic amplifiers, something of the order of 30 cycles of the a-c supply frequency. In view of the fact that magnetic amplifiers are now available having half-cycle response, it is evident that the results obtained in these early studies bear little relation to the count speeds possible with present-day types of magnetic amplifier.

The circuit shown in Fig. 2 was especially uncritical as to supply voltage variation, and could be operated over a range of ± 20 per cent. The author operated two and three successive stages of both of the circuits and in each case output signals entirely compatible with the input exigencies of the following stage were delivered. An interesting phenomenon inviting further study was noted. Under certain conditions circuit adjustments may be made whereby a trinary count action may be caused to be operative.

References

1. A Trigger Relay, W. H. Eccles, F. W. Jordan. *Radio Review*, vol. I, Oct. 1919, pp. 143–146.

2. Binary Counting Circuit, R. L. Thompson. *United States Patent 2,519,513*, Aug. 22, 1950.

3. Electrical Trigger, M. L. Wood. *United States Patent 2,524,154*, Oct. 3, 1950.

4. Saturable Reactors as Substitutes for Electron Tubes in High Speed Digital Computers, J. G. Miles. *IRE Convention Record*, New York, N. Y., March 1951.

5. Magnetic Amplifying and Control System, Alan S. FitzGerald. *United States Patent 2,027,312*, Jan. 7, 1936.

6. Feedback in Magnetic Amplifiers, A. S. FitzGerald. *Journal*, Franklin Institute, Philadelphia, Pa., vol. 247, no. 3, March 1949, and no. 5, May 1949.

7. Compensating for the Quiescent Current in Multistage Magnetic Amplifiers, Alan S. Fitzgerald. *AIEE Transactions*, vol. 71, pt. I, 1952, pp. 7–12.

On the Equivalent Circuits of Linear Amplifiers

L. M. VALLESE
ASSOCIATE MEMBER AIEE

Synopsis: The usefulness of the equivalent circuits of linear amplifiers for the circuit designer may be enhanced if the reciprocal forward- and reverse-equivalent circuits are derived. Applications to the grounded cathode vacuum-tube amplifier and to the grounded emitter transistor amplifier are shown.

THE RECENT introduction of transistors into communications has given new impetus to the study of the equivalent circuits of amplifiers. Such circuits, however, provide only a limited help for the circuit designer, because of their characteristic of nonreciprocity. As a matter of fact, for their analysis direct computation of series or parallel impedances is not permitted in general, and loop or node equations must be used. Since such equations are the basis from which the equivalent circuits are derived, these sometimes are considered as redundant.

To illustrate, consider the case of the grounded cathode vacuum-tube amplifier; see Fig. 1. Its input and output relationships are

$$\begin{cases} I_1=(G_g+j\omega C)V_1-j\omega CV_2 \\ I_2=(g_m-j\omega C)V_1+(g_p+j\omega C)V_2 \end{cases} \qquad (1)$$

The addition of external circuit elements, such as a load of admittance Y_L or a feedback branch of admittance Y_f between grid and plate, etc., corresponds to introducing additional "constraints" of the type

$$\begin{cases} I_L=-Y_LV_2 \\ I_f=Y_f(V_1-V_2),\text{ etc.} \end{cases} \qquad (2)$$

The properties of the resulting network are obtained by making simultaneous solutions of equations 1 and equation 2.

On the other hand, equations 1 may be represented in a number of ways by means of equivalent circuits, the most popular of which is shown in Fig. 1(B). When such a representation is adopted, the external circuit elements may be added directly to the equivalent circuit and the properties of the resulting network may be obtained by means of loop or node analysis. Although such a procedure may be considered to present some organizational advantages, it is not always preferred. Certainly the equivalent circuit does not provide physical interpretation of the network properties, nor does it show the effect of particular circuit elements on them.

It is the object of this paper to show that, if only some of the properties of the active networks are to be analyzed, it is possible to transform the equivalent nonreciprocal circuits into partially equivalent reciprocal ones, which have only some of their properties in common with the given amplifier. In particular, reciprocal circuits with all the forward or all the reverse characteristics can be derived.

The application of these quasi-equivalent circuits greatly simplifies the interpretation of the properties of the amplifiers and provides a direct physical representation of their properties. Furthermore, the forward and the reverse equivalent circuits might be used in combination with nonreciprocal elements such as 1-way lines[1] to synthesize amplifiers with prescribed characteristics. These "dissected amplifiers" would contain only 2-terminal elements and for this reason could be attractive for high-frequency applications.[2]

Discussion of Method

For a given linear active 4-terminal network several nonreciprocal equivalent circuit representations may be derived.[3] For example, with reference to the sign convention of Fig. 2, and assuming that the admittance matrix $\|Y_{ij}\|$ is given, and that y_{12}, y_{21}, satisfy the requirements for driving point admittances, a π equivalent circuit of the type shown in Fig. 3(A) may be used. For an arbitrary value of the admittance Y_2 the elements of the latter are[4]

$$Y_1=Y_{11}-Y_2$$
$$Y_3=Y_{22}-Y_2$$
$$I'=(Y_{12}+Y_2)V_2$$
$$I''=(Y_{21}+Y_2)V_1$$

The following particular cases are of special interest

(*a*) $Y_2=-Y_{12},\ I'=0,\ I''=(Y_{21}-Y_{12})V_1$
(*b*) $Y_2=-Y_{21},\ I'=(Y_{12}-Y_{21})V_2,\ I''=0$
(*c*) $Y_2=0,\ I'=Y_{12}V_2,\ I''=Y_{21}V_1$
(*d*) $Y_2=-(Y_{12}+Y_{21})/2,\ I'=(Y_{12}-Y_{21})V_2/2,\ I''=(Y_{21}-Y_{12})V_1/2$

Similarly, when the impedance matrix $\|Z_{ij}\|$ of the active 4-pole network is given, and z_{12}, z_{21} satisfy the requirements of driving point impedances, a T equivalent circuit of the type shown in Fig. 3(B) may be used. For an arbitrary value of the impedance Z_2 there follows

$$Z_1=Z_{11}-Z_2 \qquad Z_3=Z_{22}-Z_2$$
$$V'=(Z_{12}-Z_2)I_2 \qquad V''=(Z_{21}-Z_2)I_1$$

Paper 55-665 recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE National Electronics Conference, Chicago, Ill., October 3–5, 1955. Manuscript submitted February 14, 1955; made available for printing July 18, 1955.

L. M. Vallese is with the Polytechnic Institute of Brooklyn, Brooklyn, N. Y.

The work described in this paper has been done at the Microwave Research Institute of the Polytechnic Institute of Brooklyn, under a contract sponsored by the Office of Naval Research.

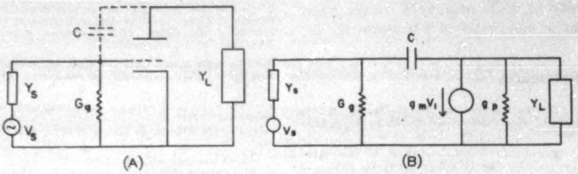


Fig. 1. Grounded cathode amplifier and its equivalent circuit

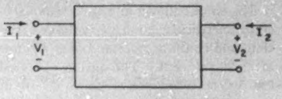


Fig. 2. Four-terminal network

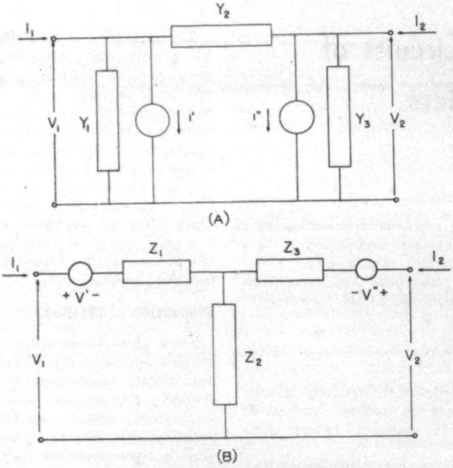


Fig. 3. Nonreciprocal equivalent circuits

A—Of π type
B—Of T type

Table I. Forward and Reverse Characteristics of Linear Active 4-Pole Networks

Characteristics	Admittance Matrix Elements	Impedance Matrix Elements
Input admittance or impedance	$Y_{11}-\dfrac{Y_{12}Y_{21}}{Y_{22}+Y_L}$	$Z_{11}-\dfrac{Z_{12}Z_{21}}{Z_{22}+Z_L}$
Forward current gain	$\dfrac{Y_{21}Y_L}{Y_{11}(Y_{22}+Y_L)-Y_{12}Y_{21}}$	$\dfrac{-Z_{21}}{Z_{22}+Z_L}$
Forward voltage gain	$\dfrac{-Y_{21}}{Y_{22}+Y_L}$	$\dfrac{Z_{21}Z_L}{Z_{11}(Z_{22}+Z_L)-Z_{12}Z_{21}}$
Forward power gain	$\left\lvert\dfrac{Y_{21}}{Y_{22}+Y_L}\right\rvert^2\dfrac{G_L}{G_{in}}$	$\left\lvert\dfrac{Z_{21}}{Z_{22}+Z_L}\right\rvert^2\dfrac{R_L}{R_{in}}$
Output admittance or impedance	$Y_{22}-\dfrac{Y_{12}Y_{21}}{Y_{11}+Y_S}$	$Z_{22}-\dfrac{Z_{12}Z_{21}}{Z_{11}+Z_S}$
Reverse current gain	$\dfrac{Y_{12}Y_S}{Y_{22}(Y_{11}+Y_S)-Y_{12}Y_{21}}$	$\dfrac{-Z_{12}}{Z_{11}+Z_S}$
Reverse voltage gain	$\dfrac{-Y_{12}}{Y_{11}+Y_S}$	$\dfrac{Z_{12}Z_S}{Z_{22}(Z_{11}+Z_S)-Z_{12}Z_{21}}$
Reverse power gain	$\left\lvert\dfrac{Y_{12}}{Y_{11}+Y_S}\right\rvert^2\dfrac{G_S}{G_{out}}$	$\left\lvert\dfrac{Z_{12}}{Z_{11}+Z_S}\right\rvert^2\dfrac{R_S}{R_{out}}$

Note: G_L, R_L are respectively the real parts of the load admittance and of the load impedance; G_{in}, R_{in} are respectively the real parts of the input admittance and of the input impedance; G_{out}, R_{out} are respectively the real parts of the output admittance and of the output impedance.

The following particular cases are of special interest

(a') $Z_2=Z_{12}$, $V'=0$, $V''=(Z_{21}-Z_{12})I_1$
(b') $Z_2=Z_{21}$, $V'=(Z_{12}-Z_{21})I_2$, $V''=0$
(c') $Z_2=0$, $V'=Z_{12}I_2$, $V''=Z_{21}I_1$
(d') $Z_2=(Z_{12}+Z_{21})/2$, $V'=(Z_{12}-Z_{21})I_2/2$, $V''=(Z_{21}-Z_{12})I_1/2$

Cases a and a' are more commonly used; cases d and d' correspond to the combination of a reciprocal network and a gyrator.[4,5]

Table I[6] summarizes the expressions of the forward and of the reverse characteristics of an active network in terms of the admittance and of the impedance matrix elements. It may be noticed that the dependence of the forward characteristics upon Y_{12} and Z_{12} and of the reverse characteristics upon Y_{21} and Z_{21} is such that it is possible to synthesize reciprocal networks with all the forward or with all the reverse characteristics of a given active network. For example a reciprocal network with admittance matrix

$$\begin{Vmatrix} Y_{11}' & Y_{21} \\ Y_{21} & Y_{22} \end{Vmatrix}$$

is forward equivalent to the nonreciprocal network of admittance matrix $\|Y_{ij}\|$ provided

$$Y_{11}'=Y_{11}+\frac{Y_{21}(Y_{21}-Y_{12})}{Y_{22}+Y_L}$$

The realization of such network is not unique; one of the possible solutions is that shown in Fig. 4(A) which is obtained from the reciprocal part of the circuit of Fig. 3(A), case B, adding an input branch of admittance $Y_{21}(Y_{21}-Y_{12})/(Y_{22}+Y_L)$. The circuit of Fig. 4(A) may contain negative elements, but no other sources.

Similarly it is seen that a reciprocal network of impedance matrix

$$\begin{Vmatrix} Z_{11}' & Z_{21} \\ Z_{21} & Z_{22} \end{Vmatrix}$$

is forward equivalent to the nonreciprocal network of impedance matrix $\|Z_{ij}\|$, provided

$$Z_{11}'=Z_{11}+\frac{Z_{21}(Z_{21}-Z_{12})}{Z_{22}+Z_L}$$

One of the possible realizations of this network is shown in Fig. 4(B); this is

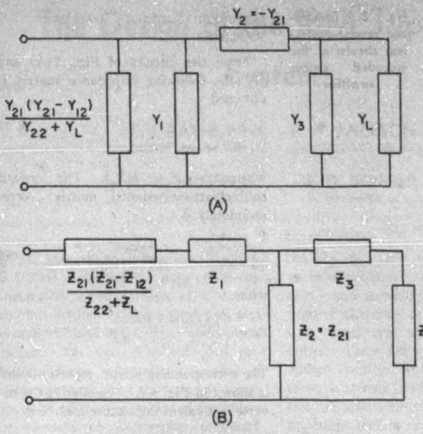

Y₂ = −Y₂₁
$\frac{Y_{21}(Y_{21}-Y_{12})}{Y_{22}+Y_L}$
Y₁
Y₃
Y_L
(A)
$\frac{Z_{21}(Z_{21}-Z_{12})}{Z_{22}+Z_L}$
Z₁
Z₃
Z₂ = Z₂₁
Z_L
(B)

Fig. 4. Reciprocal forward-equivalent circuits

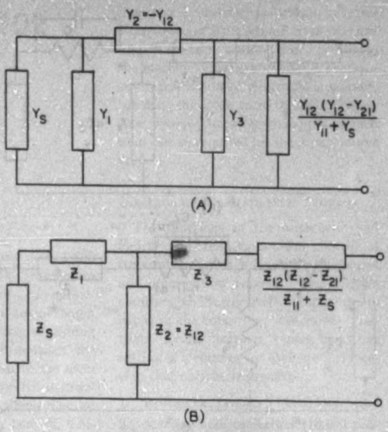

Y₂ = −Y₁₂
Y_S
Y₁
Y₃
$\frac{Y_{12}(Y_{12}-Y_{21})}{Y_{11}+Y_S}$
(A)
Z₁
Z₃
$\frac{Z_{12}(Z_{12}-Z_{21})}{Z_{11}+Z_S}$
Z_S
Z₂ = Z₁₂
(B)

Fig. 5. Reciprocal reverse-equivalent circuits

obtained from the reciprocal part of the circuit of Fig. 3(B), case *b′*, by adding a series input element of impedance $Z_{21}(Z_{21}-Z_{12})/(Z_{22}+Z_L)$.

On the other hand, a reciprocal network with admittance matrix

$$\begin{Vmatrix} Y_{11} & Y_{12} \\ Y_{12} & Y_{22}' \end{Vmatrix}$$

is reverse equivalent to the given nonreciprocal network of admittance matrix $\|Y_{ij}\|$ provided

$$Y_{22}' = Y_{22} + \frac{Y_{12}(Y_{12}-Y_{21})}{Y_{11}+Y_S}$$

One possible realization of such circuit is shown in Fig. 5(A) obtained from the reciprocal part of the Fig. 3(A) circuit case *a*, adding a shunt output branch of admittance $Y_{12}(Y_{12}-Y_{21})/Y_{11}+Y_S)$.

Finally, if the given nonreciprocal network is specified by means of its impedance matrix $\|Z_{ij}\|$, its reverse equivalent reciprocal network possesses the impedance matrix

$$\begin{Vmatrix} Z_{11} & Z_{12} \\ Z_{12} & Z_{22}' \end{Vmatrix}$$

with

$$Z_{22}' = Z_{22} + \frac{Z_{12}(Z_{12}-Z_{21})}{Z_{11}+Z_S}$$

Such a network may be realized as in Fig. 5(B), which may be considered derived from the reciprocal part of that of Fig. 3(B) case *a′*, adding a series output element of impedance $Z_{12}(Z_{12}-Z_{21})/(Z_{11}+Z_S)$.

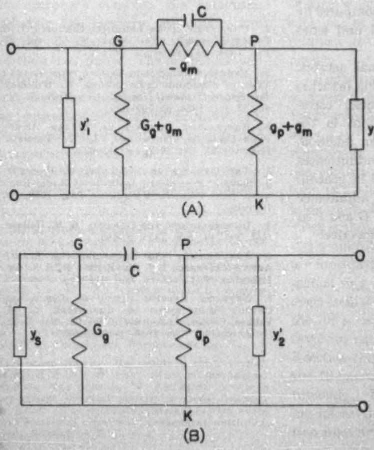

C
G
P
−gm
y′₁
Gg+gm
gp+gm
yL
K
(A)
G
C
P
yS
Gg
gp
y′₂
K
(B)

Fig. 6 (left). Forward- and reverse-equivalent circuits of the grounded cathode amplifier

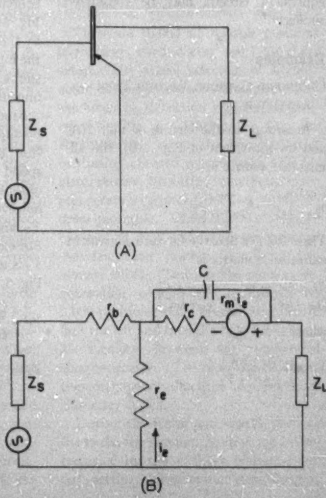

Zs
ZL
(A)
C
rb
rc
rm ie
−
+
Zs
re
ie
ZL
(B)

Fig. 7 (right). Grounded emitter transistor amplifier and its equivalent circuit

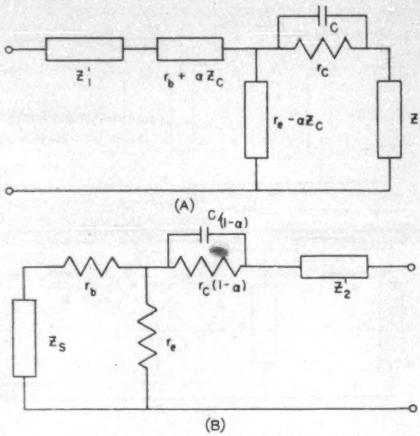


Fig. 8. Forward- and reverse-equivalent circuits of the grounded emitter amplifier

Along the lines just indicated many other quasi-equivalent reciprocal circuits may be derived. For example, if only the input admittance is of interest, the matrix

$$\begin{Vmatrix} Y_{11}'' & Y_{12} \\ Y_{12} & Y_{22} \end{Vmatrix}$$

with $Y_{11}''=Y_{11}+Y_{12}(Y_{12}-Y_{21})/(Y_{22}+Y_L)$ may be considered.

The corresponding reciprocal equivalent circuit might be made similar to that of Fig. 3(A), case *a*, with an added shunt branch. Other possible cases of quasi-equivalent circuits may be considered similarly.[7]

Examples

GROUNDED CATHODE VACUUM-TUBE AMPLIFIER

According to the circuit of Fig. 1(A) and its equivalent of Fig. 1(B), the admittance matrix is

$$\begin{Vmatrix} G_g+j\omega C & -j\omega C \\ g_m-j\omega C & g_p+j\omega C \end{Vmatrix}$$

There follows that the forward equivalent admittance matrix is

$$\begin{Vmatrix} G_g+j\omega C+Y_1' & g_m-j\omega C \\ g_m-j\omega C & g_p+j\omega C \end{Vmatrix}$$

where

$$Y_1'=\frac{g_m(g_m-j\omega C)}{g_p+j\omega C+Y_L}$$

The corresponding representative circuit is shown in Fig. 6(A). Similarly the reverse equivalent admittance matrix is

$$\begin{Vmatrix} G_g+j\omega C & -j\omega C \\ -j\omega C & g_p+j\omega C+Y_2' \end{Vmatrix}$$

where

$$Y_2'=\frac{j\omega C g_m}{G_g+j\omega C+Y_S}$$

The corresponding representative circuit is shown in Fig. 6(B).

Only elementary computations are needed for the analysis of these quasi-equivalent circuits. For example, in Fig. 6(A) the voltage gain is determined by the voltage divider network *GPK*, and the input impedance may be computed by simple evaluation of parallel and series impedances.

Modifications of the original network by addition of external shunting branches (for example, shunt feedback) correspond to identical modifications in the quasi-equivalent circuits; in addition, however, the added shunt admittances Y_1' and Y_2' must be changed in general. For example, if an element of admittance Y_3 is added in shunt between *G* and *P* in Fig. 6, Y_1' and Y_2' become respectively

$$Y_1'=\frac{g_m(g_m-j\omega C-Y_3)}{g_p+j\omega C+Y_3+Y_L}$$

$$Y_2'=\frac{g_m(j\omega C+Y_3)}{G_g+j\omega C+Y_3+Y_S}$$

GROUNDED EMITTER TRANSISTOR AMPLIFIER

From the circuits of Figs. 7(A) and (B) the following impedance matrix is obtained

$$\begin{Vmatrix} r_b+r_e, & r_e \\ r_e-\alpha z_c, & r_e+(1-\alpha)z_c \end{Vmatrix}$$

where $z_c=r_cX_c/(r_c+X_c)$. The forward equivalent impedance matrix correspondingly is

$$\begin{Vmatrix} r_b+r_e+Z_1' & r_e-\alpha z_c \\ r_e-\alpha z_c & r_e+(1-\alpha)z_c \end{Vmatrix}$$

with

$$Z_1'=\frac{\alpha z_c(\alpha z_c-r_e)}{(1-\alpha)z_c+r_e+Z_L}$$

The corresponding circuit representation is shown in Fig. 8(A). Similarly, the reverse equivalent impedance matrix is

$$\begin{Vmatrix} r_b+r_e & r_e \\ r_e & r_e+(1-\alpha)z_c+Z_2' \end{Vmatrix}$$

with

$$Z_2'=\frac{\alpha r_e z_c}{r_b+r_e+Z_S}$$

The corresponding circuit representation is shown in Fig. 8(B).

These quasi-equivalent circuits are considerably simplified when the capacitance C_c is neglected. In the general case, they prove of great help in visualizing the properties of the amplifier and in permitting simple manipulations of the various components.

References

1. UNILATERAL FOUR-TERMINAL CIRCUITS, J. S. Foley. *Electronics*, New York, N. Y., vol. 27, Feb. 1954, pp. 186–87.

2. NEGATIVE RESISTANCE ARISING FROM TRANSIT TIME IN SEMICONDUCTOR DIODES, W. Shockley. *Bell System Technical Journal*, New York, N. Y., vol. 33, July 1954, pp. 799–826.

3. EQUIVALENT CIRCUITS OF LINEAR ACTIVE FOUR-TERMINAL NETWORKS, L. C. Peterson. *Ibid.*, vol. 27, Oct. 1948, pp. 593–622.

4. THE GYRATOR AS A 3-TERMINAL ELEMENT, J. Shekel. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 42, Aug. 1953, pp. 1014–16.

5. UNDERSTANDING THE GYRATOR, L. M. Vallese. *Ibid.*, vol. 44, Apr. 1955, p. 483.

6. TERMINOLOGY AND EQUATIONS FOR LINEAR ACTIVE NETWORKS, L. J. Giacoletto. *RCA Review*, Princeton, N. J., vol. 14, Mar. 1953, pp. 28–46.

7. NETWORK REPRESENTATION OF INPUT AND OUTPUT ADMITTANCES OF AMPLIFIERS, L. M. Vallese. *American Journal of Physics*, New York, N. Y., vol. 17, Nov. 1949, pp. 482–86.

Vol
Pt.
195

The Design of the IBM Type 702 System

C. J. BASHE
ASSOCIATE MEMBER AIEE

P. W. JACKSON
ASSOCIATE MEMBER AIEE

H. A. MUSSELL
NONMEMBER AIEE

W. D. WINGER
NONMEMBER AIEE

ol. 74
. 1
955

THE International Business Machine (IBM) electronic data-processing machine type 702 consists of a stored-program arithmetic and logical unit and an array of high-speed input and output devices. The system is intended primarily for the automatic processing of business records, and its design is therefore quite different from that of computers for scientific or engineering problems. Some of the more important features are:

1. Electrostatic memory capacity of 10,000 alphabetic or numeric characters.
2. Variable word and record (or input-output message) length.
3. A wide range of input and output devices, including magnetic tapes, punched cards, keyboard, typewriter, high-speed printer, and auxiliary storage in the form of magnetic drums, which are treated logically as an input-output medium; see Fig. 1.
4. A choice of interconnection of machine units as indicated in Fig. 2.

Arithmetic and Logical Unit

The arithmetic and logical unit consists of three separate machine frames: the operator's console; the computer, including the memory, arithmetic, and control circuits; and the power supply for these two units. The arithmetic and control circuits, which will be described first, perform the functions of basic timing, control of electrostatic memory, program control, and all arithmetic and manipulative operations.

The arithmetic and control circuits occupy about half of the computer cabinet; see Fig. 2. They comprise over 1,600 vacuum tubes, about 4,300 germanium diodes, and associated circuits. The functions of this part of the machine begin with the generation of clock or timing pulses. These are used to advance the address counters which control the electrostatic storage and to step the instruction and execution timers which play an important part in program control. Finally, the arithmetic element and the manner in which various instructions are carried out will also be discussed in this section.

TIMING

The requirements of the timing (or clock) circuits are best described by reference to Fig. 3, which represents the basic machine cycle, or "character cycle." Since the timing is governed primarily by the requirements of electrostatic storage, the terminology used in connection with major portions of the cycle is that of electrostatic storage. In general, characters are read from electrostatic storage during the READ portion of the cycle, operated upon during the ADD portion (during which one bit is regenerated in each pair of storage tubes) and the result returned to storage during the WRITE portion.

All information is transmitted as groups of seven parallel binary digits (or bits). Each seven-bit group represents a character, and characters are transmitted sequentially (or serially) with respect to one another. An example of the coded form of a character, as written on paper is shown in Table I. The character code chart is shown in Appendix I.

The timing signals for a character cycle are furnished by a clock consisting of a 3-section ring circuit driven by a 1-megacycle oscillator, as shown in Fig. 4. When the 702 is performing any instruction not involving an input or output device, the 3-section ring closes on itself and caused the 23-microsecond cycle to repeat continuously. If, however, the 702 is sending information to or receiving information from an input or output device, the character cycle ring is broken and the 9-microsecond REPEATED REGENERATION ring is closed. Under these conditions, the machine remains normally in continuous regeneration of electrostatic storage, breaking this cycle to go through a READ, ADD, WRITE character cycle only at such intervals as required by the input-output unit.

CONTROL OF ELECTROSTATIC STORAGE

The functions of the arithmetic and logical unit relating to the control of electrostatic storage are address control and information control. The address counters and associated equipment are listed in the following; the word "positions" means "settings," there being one setting of a counter for each storage location that may be addressed.

1. Instruction counter (10,000 positions).
2. Memory address register (10,000 positions).
3. Memory address counter (10,000 positions).
4. Memory regeneration counter (2,000 positions).
5. Memory deflection register (10,000 positions).
6. Starting-point counter *A* (512 positions).
7. Starting-point counter *B* (512 positions).
8. Accumulator address counter (512 positions).
9. Accumulator regeneration counter (1,024 positions).
10. Accumulator deflection register (512 positions).

There are 10,000 addressable positions in memory, each storing one 7-bit character. The actual selection of a character address is always accomplished by the memory deflection register, as indicated in Fig. 5. During the ADD (or REGENERATE) part of the character cycle, or during repeated regeneration, this register receives its setting from the memory regeneration counter, which steps one address per cycle. Physically the memory consists of five sets of seven memory drawers, each drawer containing two storage tubes. Thus, as the memory regeneration counter steps through its 2,000 positions, the corresponding bits of first the left and then the right tubes of all 35 memory drawers are regenerated simultaneously. The complete regeneration of memory therefore requires 2,000 character cycles.

During the READ and WRITE parts of the cycle, the memory position used is determined by the memory address counter's setting. This counter steps along by

---

Paper 55-719, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted May 19, 1955; made available for printing July 20, 1955.

C. J. BASHE, P. W. JACKSON, H. A. MUSSELL, and W. D. WINGER are with the International Business Machines Corporation, Poughkeepsie, N. Y.

This paper describes the results of the concerted efforts of many engineers. Much of the original planning was done by N. Rochester, in whose department the final engineering model was also produced. The electrostatic storage system was a modification of a previous development which had been directed by P. E. Fox.

Fig. 1. View of 702 installation

one character address per cycle as the seven bits of each character of a word are read from seven cathode-ray tubes simultaneously. The initial setting of the counter is taken from the instruction counter or the memory address register, depending on whether an instruction is being read from memory or is being executed. The memory address counter must be able to step up or down through the memory, because arithmetic processes require handling from lowest-to-highest-order digits, while input-output functions require the reverse.

There are two 512-position electrostatic accumulators, referred to as accumulator *A* and accumulator *B*, which perform the function of holding arithmetic subtotals, partial products, multipliers, and quotients, as well as large blocks of data to be transferred from one location in memory to another. These two accumulators are used in exactly the same manner, and a "tag" on the instruction specifies which is to be used. The address in the accumulator is never referred to in the program, and addressing may therefore be thought of as relative to the setting of the starting-point counter. The block diagram of accumulator address control is represented in Fig. 6. Straight binary addressing is used for economy of counting; since the programmer need not be concerned with the address, the binary counters introduce no inconvenience.

The flow of information to and from electrostatic storage is shown in Fig. 7, where the solid lines denote paths during the READ part of the cycle and broken lines denote paths taken during WRITE time. This diagram will be referred to in the section on arithmetic functions.

PROGRAM CONTROL

The IBM 702 is a single-address stored-program machine. Each instruction of the program consists of an operation part and an address part; the former identifying the nature of the operation to be performed, and the latter specifying for most instructions the location in memory of the data involved in the operation. The 702 is capable of performing 32 kinds of operations, and therefore one character is sufficient to designate the operation (since 64 combinations of the 4-digit bits and the two zone bits of a character are possible).

The location of a word is specified by the address in memory of its right-hand (or lowest-order) character. The address part of the instruction usually specifies the location of the word to be operated upon; a word can be of any length, and the address part of the instruction must therefore be able to specify any one of 10,000 memory positions. Thus a 4-digit address part is necessary for each instruction.

One further feature of the instruction is that the thousands digit of the address part may be tagged if accumulator *B* is to be used during the operation, whereas accumulator *A* would otherwise be used. Tagging the thousands digit is accomplished by imparting "negative" zoning to the character, i.e., giving its *B* and *A* bits the binary combination "10" whereas they would normally have the value "00" used for purely numeric characters. The existence of two accumulators is incidental to a basic description of the 702 and will not be discussed in any greater detail.

A typical instruction, then, might appear in memory as *G6127*, where *G* denotes the operation ADD. (This list of

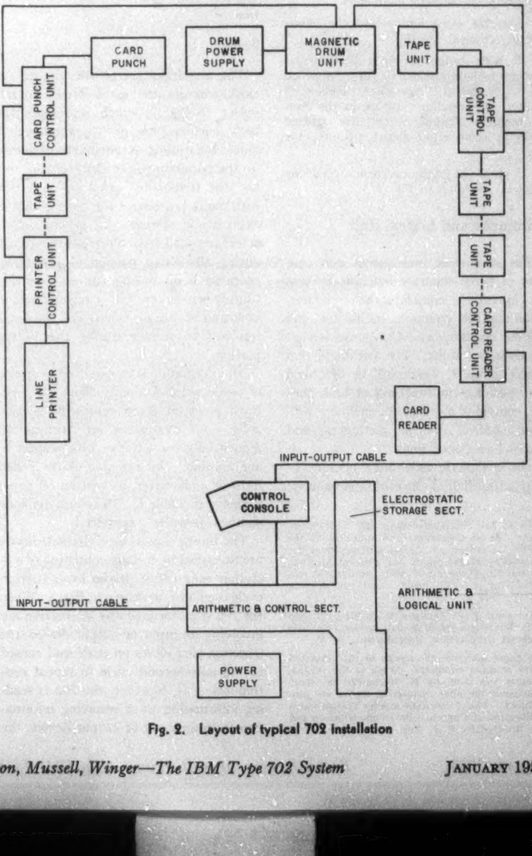


Fig. 2. Layout of typical 702 installation

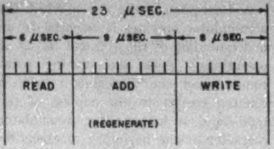
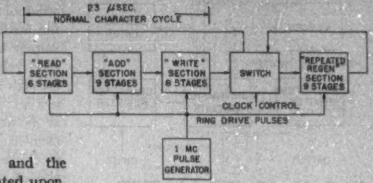

23 µSEC.
6 µSEC.
9 µSEC.
8 µSEC.
READ
ADD
WRITE
(REGENERATE)

Fig. 3 (left). Character cycle of the 702

Fig. 4 (right). Block diagram of clock circuits


23 µSEC.
NORMAL CHARACTER CYCLE
"READ" SECTION 6 STAGES
"ADD" SECTION 9 STAGES
"WRITE" SECTION 8 STAGES
SWITCH
"REPEATED REGEN" SECTION 9 STAGES
CLOCK CONTROL
RING DRIVE PULSES
1 MC PULSE GENERATOR

operations is given in Appendix II.) The 702 would interpret this instruction as meaning "Add to the contents of the accumulator the word whose lowest order digit is stored at address 6127 in memory."

Instructions are stored in memory, generally in continuous blocks of addresses, and are read and executed in sequence except where transfers of control break the sequence.

The manner of operation of the instruction counter should be mentioned here. This counter provides only the units digit address setting for each instruction and, after the memory address counter is set to this same address, the instruction counter steps upward to the units digit of the next instruction. Meanwhile, the memory address counter, is stepping downward through successive characters of the instruction being read from memory. This seemingly elaborate procedure allows the instruction to be read units digit first, and the program to be stored with successive instructions in higher-addressed memory locations, while not requiring the instruction counter to count both up and down. The desirability of reading an instruction units digit first stems from the fact that the address of this digit is used as the address or location in memory of the 5-character instruction word; hence, a transfer of control to a particular instruction is specified as a transfer to the location of the units digit of the instruction. The operation of the counters may be clarified by reference to Fig. 8, showing the instruction P7392 being read from memory.

Instructions are read from memory during instruction time and executed during execution time. These two periods alternate regularly. Each instruction time lasts for the 5-character cycles required to read the instruction, and execution time lasts for a variable number of cycles depending upon the operation and the length of the word being operated upon. The instruction timer and a typical execution timer (for the operation ADD) are represented in Fig. 9. The instruction timer is a simple ring of five vacuum-tube trigger (bistable) stages, receiving one "advance" pulse per character cycle. The execution timer consists of similar bistable stages with more complex gating on each stage. The execution timer is conditioned at the end of each character cycle, according to the manner in which the operation is progressing, either to initiate the next type of cycle and end the current type or to continue in the current type of cycle for another character cycle.

Table I. Example of Coded Form of Character

	Check Bit	Zone Bits		Digit Bits				
Bit Name	C	B	A	8	4	2	1	
Value	1	0	0	0	1	1	1	=7 (decimal)

The numbers on the blocks denoting the type of the cycles during execution time should be distinguished from those denoting distinct single cycles during instruction time. Incidentally, the flow lines shown leaving and entering the same type of cycle during execution time merely indicate that a trigger stage is left on, rather then representing a dynamic-delay mode of storage, since all control triggers, as well as register and counter stages, are bistable twin-triode circuits. It is the execution timers, more than any other part of the machine, in which the unique features of control required by a variable-word-length computer are evident.

Closely associated with program control are the memory address register, the operation register, and the operation decoder (see Fig. 10). The memory address register is a 4-decimal-order register which receives, during the first 4 cycles of instruction time, the four digits of the address part of the instruction read from memory. The setting of this register is transferred, during the fifth cycle of instruction time, to the memory address counter, see Fig. 5, and designates the position in memory of the first character of the word to be operated upon during execution time.

During the fifth character cycle of instruction time, the operation part of the instruction is read from memory and stored in the operation register. The six bits exclusive of the check bit are fed as bipolar signal voltages into the germanium-diode-matrix operation decoder. With six pairs of input wires, there are 64 possible combinations. As it happens, 32 of these are needed, and these appear as individual control wires at the output. Several checks are made upon the correctness of an instruction, and only two of

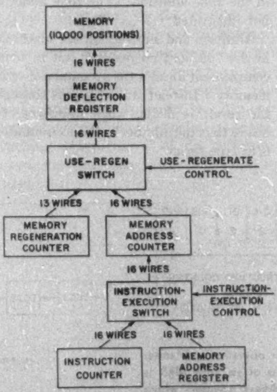

MEMORY (10,000 POSITIONS)
16 WIRES
MEMORY DEFLECTION REGISTER
16 WIRES
USE-REGEN SWITCH
USE-REGENERATE CONTROL
13 WIRES
16 WIRES
MEMORY REGENERATION COUNTER
MEMORY ADDRESS COUNTER
16 WIRES
INSTRUCTION-EXECUTION SWITCH
INSTRUCTION-EXECUTION CONTROL
16 WIRES
16 WIRES
INSTRUCTION COUNTER
MEMORY ADDRESS REGISTER

Fig. 5. Memory address control

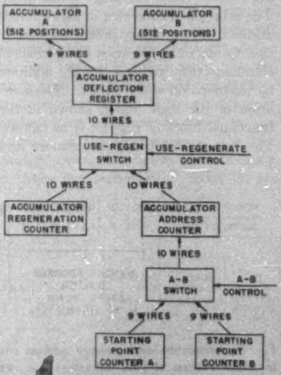

ACCUMULATOR A (512 POSITIONS)
ACCUMULATOR B (512 POSITIONS)
9 WIRES
9 WIRES
ACCUMULATOR DEFLECTION REGISTER
10 WIRES
USE-REGEN SWITCH
USE-REGENERATE CONTROL
10 WIRES
10 WIRES
ACCUMULATOR REGENERATION COUNTER
ACCUMULATOR ADDRESS COUNTER
10 WIRES
A-B SWITCH
A-B CONTROL
9 WIRES
9 WIRES
STARTING POINT COUNTER A
STARTING POINT COUNTER B

Fig. 6. Accumulator address control

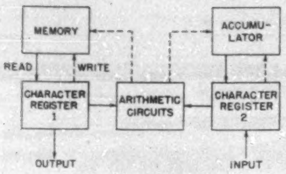


Fig. 7. Information flow during READ and WRITE portions of character cycle

these will be remarked upon here, since checking is described in a later section. The unused possible outputs of the operation decoder are all recognized as error signals, and serve to indicate a programmer's error or the fact that instructions are not being read correctly from memory. Another check, shown in Fig. 10, is made possible by the fact that the correct *C*-bit for each output is predetermined and the outputs may simply be fed into OR circuits for comparison with the status of the *C* trigger of the operation register. This is a check on the correctness of the operation part as well as its interpretation.

ARITHMETIC FUNCTIONS

The block in Fig. 7 labeled "Arithmetic Circuits" includes the adder and special circuits for complementing, multiplication, and comparison. The adder, the logical center of all these circuits and involved in every arithmetic operation, will be described first.

Because the bits of each decimal digit are transmitted in parallel in the 702, it has been possible to employ an adding circuit which is both simple and free of high-speed pulse-handling components. The principle used is that of logical diode addition of four binary orders, the correction for decimal operation being effected simultaneously with the binary addition by modification of the circuit. The operation of the digit adder, shown in the functional block of Fig. 11, is as follows. The four digit bits of the memory and accumulator characters are presented to the arithmetic circuits as direct-coupled voltage by character register 1 and 2 respectively. The accumulator digit passes through the true-complement unit, which is a set of switching circuits, and arrives at the adder in either true or nines-complement form. If the true-complement unit is set to COMPLEMENT (as in subtraction of positive numbers) a decimal in-carry is supplied to the digit adder to effect tens-complementing.

The 4-order binary adder needs no explanation, beyond the fact that it consists of conventional germanium-diode logical circuits. The decimal out-carry is generated by either a carry out of the 8-bit position or the presence of an 8-bit and a 4- or 2-bit in the uncorrected sum. This out-carry is stored and becomes the in-carry to the adder on the next cycle. The decimal out-carry is used also to control modification of the 4-order binary sum to obtain the corrected decimal sum. If a carry exists, 6 is, in effect, added; otherwise, the sum is left unaltered.

In ordinary arithmetic operations, the zone bits *A* and *B* of the operands are disregarded, and the zone bits of the sum or product therefore becomes zeros, or numeric zoning. With the ADD MEM instruction, however, it is possible for the zone bits to take part. In this operation, the zone bits *A* and *B* of each character from the memory and accumulator enter a 2-order binary adder, binary carries being propagated from the zone addition of one character to the next. The digit parts are added independently, as previously described, and decimal carries are propagated from one digit addition to the next, independently of the binary carries in the zone addition. The zone adder is not illustrated.

Addition and subtraction are algebraic in the 702, and no result is left in tens-complement form in the accumulator or in memory. Instead, an extra pass through the result is made, when necessary, to insure that the number in the accumulator is in true form.

A description of the way in which the accumulator operates is essential to an understanding of the 702 and its set of operations. Both the starting point counter and the accumulator address counter are set to the address of the right-hand or lowest-order accumulator character at the beginning of execution of an instruction; they will have been left in that position at the end of the last instruction involving the accumulator. One position to the left of the highest-order character there will always be a special character called an "accumulator mark," which serves to define the left end of the accumulator word.

If the operation is ADD, for example, the accumulator address counter advances through the accumulator to the left, while the memory address counter does the same in the memory field specified by the address part of the instruction. The result field is as long as the longer of the two numbers being added; since the result remains in the accumulator, the accumulator mark must be moved to the left if the memory field exceeds in length the original accumulator field. After the new accumulator mark is written, the accumulator address counter is set to the starting point counter and at this point the execution of the instruction is complete.

A unique feature, resulting from the variable word size and the accumulator addressing system, is the method in which the equivalent of "shifting" is accomplished. If an addition is to start with the lowest-order digit in the accumulator being added to the third digit from the right in the memory word, for example, it is necessary first to annex two zeros to the right end of the accumulator word. This is done by giving the LENG 0002 instruction, which actually increases the length of the accumulator word by two zeros at the right. The SHORTEN, ROUND, and SET LEFT operations also affect the length of the word in the accumulator in order to give the desired decimal weight to the right-hand digit.

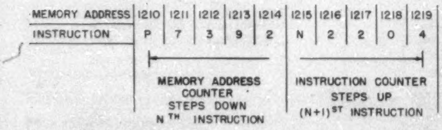

MEMORY ADDRESS
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
INSTRUCTION
P
7
3
9
2
N
2
2
0
4
MEMORY ADDRESS COUNTER STEPS DOWN N TH INSTRUCTION
INSTRUCTION COUNTER STEPS UP (N+1) ST INSTRUCTION

Fig. 8. Instruction and memory address counter operation. Manner in which address of Nth instruction, i.e., address of its units digit is imparted by instruction counter to memory address counter, which controls reading of this instruction while instruction counter steps up to address of (N+1)st instruction

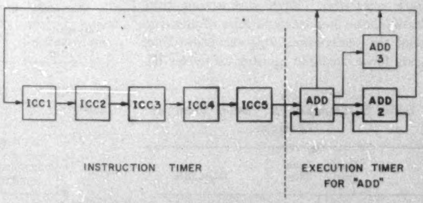


Fig. 9. The instruction timer and a typical execution timer

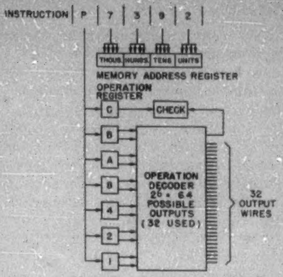


Fig. 10. Instruction interpreting equipment

The 512-character capacity of the accumulator (actually 511 characters plus accumulator mark) is of value in handling several words or a whole record with one instruction. The LOAD instruction, for example, causes a section of the memory contents of length equal to the accumulator field length to be loaded into the accumulator. This instruction is also necessary if single alphabetic words are to be taken into the accumulator, because the zoning of memory characters to indicate field boundaries, as employed in arithmetic instructions, is not feasible with nonnumerical characters.

The large capacity of the accumulator also comes into play during multiplication and division. In multiplication, the starting point counter and accumulator address counter can keep track of two words (not to be referred to simultaneously) by stepping halfway around the closed loop of 512 positions. This is done in a single step by simply inverting the highest-order bit of the binary address setting in the counter. Thus the accumulator may be regarded as a circular array of 512 1-character storage positions, and the multiplier and product are stored at diametrically opposite positions. Since the product may occupy 256 positions, minus one position for the accumulator mark, it is possible to multiply very large fields together in the 702.

Multiplication is of particular interest, in that special circuits are provided for executing this operation. The address part of the multiply (MPY) instruction specifies the location of the number in memory which is the multiplicand. The number in the accumulator is the multiplier. The process of multiplying consists of adding to the partial product, for each digit of the multiplier, the corresponding multiple of the entire multiplicand number, the latter being shifted one decimal position to the left each time. The partial product is built up in a part of the accumulator remote from the part occupied by the multiplier, and when the final product has been built up, the starting point counter is set to the address of the product, disregarding the multiplier thereafter. The fact that there are, in effect, two sections in each accumulator, should be distinguished from the fact that there are two entirely separate accumulators.

The method of obtaining the proper multiple of the multiplicand as it is fed digit by digit into the adder is illustrated in Fig. 12. Selection of multiples is controlled by the multiple register, which is set to the value of the multiplier (accumulator) digit. For each such setting, the entire multiplicand is passed through the multiple generators and multiple selectors and added to the partial product, the latter being replaced digit by digit as the new partial product is created. Note that the two numbers in the accumulator need never be read simultaneously. First the multiple register is set to the multiplier digit, and then the partial product is added to the "*N*-tupled" multiplicand. The multiple adder and its correction circuits, incidentally, are exactly like the digit adder and correction circuits of Fig. 11.

The time required for multiplication is somewhat longer than the time required for the process just described. Certain character cycles are spent in placing special marks in the accumulator to indicate the ends of fields and in stepping the accumulator address counter through the product field so that it may be left at the lowest-order digit of the product when the operation of multiplication has been ended.

This discussion should enable the reader to understand the general manner of execution of the various operations by the 702. A list of the operations, together with pertinent descriptive remarks, appears in Appendix II. A more detailed description of the behavior of the machine under control of the different operations has been given in another paper.[1]

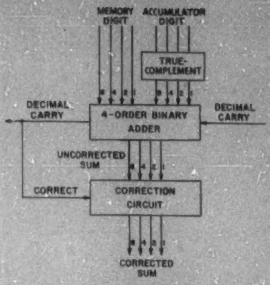


Fig. 11. Digit adder in the 702

Electrostatic Storage System

The electrostatic storage system of the 702, as viewed from the cathode-ray tube side, is shown in Fig. 13. Each removable drawer contains two cathode-ray tubes of the 3-inch diameter. *IBM-85* type, as well as a video amplifier and certain vacuum-tube circuits for signal gating and for the control of regeneration. The manner in which the 44 storage-drawer positions are allotted to memory, the accumulator, and the spare storage drawers is made clear by Fig. 14. No attempt will be made here to describe the engineering development of cathode-ray tube (CRT) memory, which is covered elsewhere in the literature.[2,3] This paper limits itself to the adaptation of the CRT memory to the 702.

When a memory group of seven drawers is selected by the three highest-order bits of the thousands digit in the deflection register, all seven drawers (i.e., all seven bits of each character) are referred to simultaneously during the READ portion of each character cycle. Similarly, all seven drawers of the selected accumulator are read from at the same time. Nearly every type of operation involving electrostatic storage results in the changing or replacing of information stored in the memory or the accumulator. The unchanged operand, whether from memory or from the accumulator, must be re-

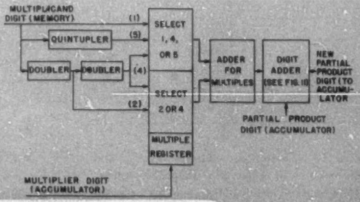

MULTIPLICAND DIGIT (MEMORY)
(1)
(5)
QUINTUPLER
DOUBLER
DOUBLER
(4)
(2)
SELECT 1, 4, OR 5
SELECT 2 OR 4
MULTIPLE REGISTER
ADDER FOR MULTIPLES
DIGIT ADDER (SEE FIG. 11)
NEW PARTIAL PRODUCT DIGIT (TO ACCUMULATOR)
PARTIAL PRODUCT DIGIT (ACCUMULATOR)
MULTIPLIER DIGIT (ACCUMULATOR)

Fig. 12. Block diagram of multiplying circuits

Fig. 13. Electrostatic storage section

turned to storage during the WRITE part of the cycle, the reason for this being that the read-out process is a destructive one.

The regeneration of information in electrostatic storage occurs at one address in all storage drawers simultaneously during the REGENERATE, or ADD, portion of each character cycle. Regeneration proceeds sequentially through all the addresses in one cathode-ray tube of each drawer and then the other, under control of the regeneration counter. Thus there is no definite relationship between the position in use and the position regenerated during a cycle. A complete regeneration of memory requires 2,000 cycles, since there are 1,000 positions in each of the two cathode-ray tubes of a drawer.

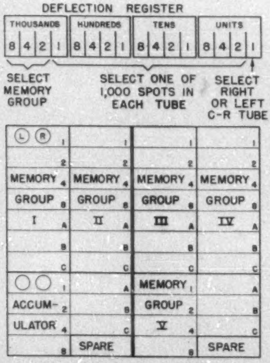


Fig. 14. Electrostatic storage layout, and the manner in which the deflection register controls selection of the memory address

The raster consists of 40 horizontal lines, with 25 bits per line. Every second line is displaced one-half a bit-space to the right, giving hexagonal packing, in which each spot on the tube has six nearest neighbors. The *IBM-85* cathode-ray tube, thus operated, gives a wide margin of safety against spill, at the maximum number of references to a given spot per complete regeneration. Address scattering is employed, so as to maximize the number of addresses which must be referred to in sequence before two adjacent spots have been used. Briefly, this is accomplished by allowing the lower-order positions of the deflection register to generate the larger increments of cathode-ray deflection.

Input-Output System

The flexibility of the 702 will become more evident through a description of the input-output system. All machine units of the installation are connected by means of a single cable, as indicated in Fig. 2. This cable contains address, control, and information handling wires.

Each type of input-output device has a control unit whose primary function is to adapt its input or output unit to the system. Thus, the logic of executing input-output instructions is independent of the type of unit selected. Each control unit supplies d-c power for the input-output units attached to it. It also contains a test station from which its associated units may be tested. In this way there is no interference with the remainder of the system.

The control units of the card reader, card punch, and printer may be detached from the arithmetic and logical unit and used to control a tape unit to provide card-to-tape, tape-to-card, and tape-to-printer operations.

The input or output device to be used in a READ or WRITE operation is specified in the address part of the last previous SELECT instruction. Since there are four digits in the address part of the instruction, it is possible to select a total of 10,000 input-output addresses. Table II illustrates the distribution of these addresses among the input-output units.

The address designated by the address part of the SELECT instruction is placed in the select register which is in the arithmetic and logical unit. It is retained there until another SELECT instruction is executed. Thus, only one address may be selected at any time. This register powers the 16 select wires in the input-output cable. Each control unit recognizes the select address assigned to it.

READING AND WRITING

The processes of reading or writing with an input-output unit are carried out asynchronously with respect to the character-cycle timing of the arithmetic and logical unit. Information is transmitted over the read and write busses in a parallel bit, serial-character fashion, with the repetition rate controlled by the input-output device in operation. The wires in the input-output cable used during reading or writing are as follows:

THE READ BUS

A group of seven wires which by their status indicate the contents of a character register containing the current input character. Transmission of information is accomplished by sampling these lines within the arithmetic and logical unit.

THE WRITE BUS

A group of seven wires which carry output characters from the arithmetic and logical unit during writing operation. Transmission of information is accomplished by sampling these lines within the output unit.

THE READ CALL LINE

A direct-coupled line which indicates by its status that a READ operation is to be executed.

THE WRITE CALL LINE

A line which indicates by its status that a WRITE operation is to be executed.

THE RESPONSE LINE

A pulsed line which is the key to the entire asynchronous operation of reading or writing. Each time an input device has set up a character to be sampled in the arithmetic and logical unit, a pulse is transmitted by the input unit indicating

Table II. Distribution of Addresses Among Input-Output Units

Address Part of Select Instruction	Input-Output Devices Selected
0100–0199	card readers*
0200–0299	tape units*
0300–0399	card punches
0400–0499	line printers*
0500–0599	typewriter
0900–0905	check indicator†
0911–0919	console switches for program alteration†
1000–9999	drum sections*

* Selection of any of these units also selects the corresponding input-output indicator as the signal for interrogation by the TRANSFER ON SIGNAL instruction; see Appendix II.

† These are selected only for interrogation by the TRANSFER ON SIGNAL instruction.

that the clock in the arithmetic and logical unit is to abandon temporarily the repeated regeneration of electrostatic memory and go through one character cycle in order to store the input character in its proper address in memory. Similarly, when an output device has recorded the current character on the write bus, it sends a pulse on the response line requesting the next character.

The Disconnect Line

A line which is pulsed by the input-output unit to indicate to the arithmetic and logical unit that the execution of the input or output operation is completed.

Control Functions

Various operations such as REWIND TAPE and TURN ON INPUT-OUTPUT INDICATOR are necessary to operate the input-output system. They are accomplished by a CONTROL instruction. The address part of this instruction designates its function. Only the units order of the address is decoded. Only five of the ten available control lines are assigned functions.

Signal Lines

The remaining lines in the input-output cable are signal lines which carry checking information or indicate that the selected input-output indicator is on.

Characteristics of Input and Output Equipment

While the design of the 702 system is such as to take advantage of the high speed of magnetic coated plastic tape as the principal direct input-output medium, other devices such as punched card readers and line printers play an important part in the system, both as direct input-output and as tape-operated auxiliary units. A brief description follows of the input-output units at present available.

Magnetic Tape Units

Unlike the rest of the input-output equipment, the tape units do not require individual control units to adapt them to the system. Instead, each magnetic tape control unit is provided with the capacity to handle up to 10 tape units. Hence the control unit recognizes as its own a group of 10 SELECT addresses and examines the units digit of the address to determine which to activate.

The $10\frac{1}{2}$-inch-diameter reels of the tape unit accommodate 2,400 feet of 1/2-inch-wide magnetic oxide-coated plastic tape. The tape travels at a rate of 75 inches per second during READ, WRITE, and BACKSPACE operations. Acceleration of the tape from a standstill requires 10 milliseconds. A gap of 3/4 inch is allowed between successive records. Reflecting strips of adhesive tape are placed near the beginning and end of each tape. These are sensed photoelectrically to indicate the starting point (load/rewind point) and the end-of-tape respectively.

To achieve fast acceleration the tape unit employs storage of slack tape from each reel in vacuum columns. The magnetic read-write head and photoelectric pickups, along with the drive capstans and the actuating mechanism, are between the two vacuum columns. Pressure-sensitive switches in the vacuum columns provide signals to the reel-driving clutches to maintain the proper amount of slack tape.

A high-speed rewind allows a full reel to be rewound in 1.2 minutes. The tape unit automatically pulls the tape out of the vacuum columns, opens the magnetic head housing, and retracts the drive capstans so that a direct reel-to-reel transfer of tape can be effected. The tape is automatically reloaded at the end of the rewind sequence. The automatic loading feature is useful when placing tape on the machine in that the operator need only attach the tape to the empty reel. The remainder of the loading process is automatic.

Records may be of any length within the memory capacity of 10,000 characters. The end of a record is determined when reading by sensing the inter-record gap. In writing, the end of a record is signaled by a "record mark" in the memory unit. The tape may be backspaced one record at a time by the CONTROL 0004 instruction.

A nonreturn-to-zero method of recording is used.[4] In this method the recording head always saturates the tape in one polarity or the other. A "1" is recorded by reversing the current in the appropriate write head so that the polarity of the saturation recording on tape reverses at that point.

A "0" is recorded as no change in flux, and therefore as continued magnetic saturation in the same direction. On reading, the sensing of a change in flux in any track indicates a "1" in that track. The presence of a character is announced by the sensing of a "1" in any of the seven tracks, the code being so arranged, Appendix I, that every character has at least two "1s." Recording is parallel by bit, serial by character. The density is 200 characters per inch.

Card Readers and Card Punches

The card reader is designed around a parallel card feed similar to that which is at present used in standard IBM accounting machines. This feed furnishes card information to an 80-character magnetic core buffer storage unit at the rate of 250 cards per minute.

The magnetic core buffer storage makes it possible to transfer a punched-card record to the computer at a high character rate, and allows the computer to continue the program while the information from the next card is recorded in the buffer. Since cards are punched in the 12-hole IBM code, the buffer was designed to store a complete card image and it consists of a 2-dimensional 12 by 80 matrix. Information enters the buffer parallel by character and serial by bit, and is transferred to the computer serial by character and parallel by bit after appropriate translation.

The card punch is built around a standard, 100-card-per-minute punch unit. A 12 by 80 core buffer storage unit is included, to allow high-speed information transfer from the computer to the punch without slowing down the program.

Printer

The line printer is a wheel-type printer which prints 120 characters per line at a speed of 150 lines per minute. A buffer consisting of a 12-by-120-core matrix stores the information for printing.

A special feature of the printer operation is that the first character of a record may be used to control carriage operation. This operation takes place before the information accompanying this character is printed. The line printer and control unit may be detached from the 702 system and may be operated in conjunction with a tape unit for tape-to-printer operation.

Typewriter

An electric typewriter output unit is incorporated in the operator's console, and is available for use under program control as an auxiliary printer to print miscellaneous operating data. Program testing is a typical example of use of the typewriter. No buffer is provided with the typewriter, and the program must therefore wait until an entire record is typewritten before it can continue. The rate of typing is 10 characters per second.

Magnetic Drums

The addressing system provides addresses 1,000 to 9,999 for sections of magnetic drums. Each drum is capable

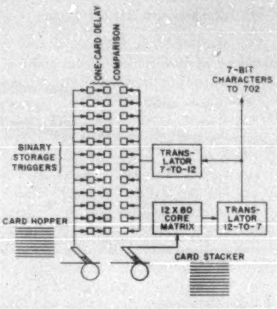


Fig. 15. The method by which the card reader is checked. Three groups of 12 trigger circuits store the row counts (odd-even) for a full card

of holding 60,000 characters (six times the capacity of the electrostatic memory) in a total of 3,000 sections of 200 characters each. The sections are addressed individually, but a single continuous record may extend over a number of sections, limited only by the total capacity of the drum. More than one drum may be provided, however.

The drum, a brass cylinder about 10.7 inches in diameter, is plated with a 20-per-cent nickel 80-per-cent-cobalt surface 1/2 mil thick. Recording is at a density of 60 characters to the inch, each character being recorded on seven parallel tracks. Successive characters are interleaved in such a way that only every fifth character encountered by the read-write heads (in reading, for example) appears on the read bus. The resulting rate of information transfer is 25,000 characters per second.

LINE AMPLIFIERS

The system as described will operate with a total of 10 input-output control units (including the console typewriter). Since each magnetic tape control unit will serve 10 tape units, this should be a rather large installation. If more than 10 control units are required, one or more line amplifiers must be added to the system, with each unit acting as a signal repeater for up to 10 additional control units.

Checking

The 702 is provided with a generous complement of circuits for checking its correct operation. Specifically, the following functions are checked automatically:

1. Reading of instructions from the internal memory, and their correct interpretation, as described previously in the section entitled "Program Control."
2. Transmission of information from the internal memory.
3. Reading of data from punched cards, magnetic tapes or drums, and transmission of the data to the arithmetic and logical unit, transmission of data from the arithmetic and logical unit to tapes, drums, card punches, and printers.
4. Correct actuation of the print magnets in the line printer, and correct operation of the card punch.
5. Correct-result word length assignment in arithmetic operations, by means of an "overflow check."
6. Correct addressing of words in arithmetic operations, by means of a "sign check."

Two of the more significant checks will now be described in greater detail.

CHECKING CARD READER

The card reader contains two brush stations, the first of which is used for obtaining a row-by-row check sum (modulo 2) and the second for actually reading the card (row by row) into the 12-by-80 magnetic core card buffer. Translation from the 12-bit card code to the 7-bit internal 702 code is accomplished during the serial (column by column) read-out from the buffer to the 702, and it is during this time that the check is performed, as shown in Fig. 15.

Each character, as it is transmitted to the 702, is also retranslated from 7-bit to 12-bit code for the purpose of developing a modulo 2 check sum for each row, which can be compared against the check sum obtained at the first reading station. Since both reading and translation are thereby checked, and the redundancy check in the 7-bit code detects any errors in transmission to the electrostatic memory in the 702, the entire input operation is checked.

CHECKING MAGNETIC TAPE

The recording of each character on magnetic tape is checked by returning to the arithmetic and logical unit the 7-bit "echo" character obtained by sensing the flow of current in the seven write heads, and performing the redundancy check upon it. Any error so indicated may be detected by the program at the end of the WRITE operation so that the tape may be backspaced and the writing process repeated while the information is still available within the memory.

A double check on reading from tape is provided. During the writing process, a "check character" is written at the end of the record, just three character spaces after the last character. The 3-character gap identifies the end of the record, and gives notice that the check character is about to be read. The check character is such as to cause the total count of "1s" in each of the seven tape track i.e., the "longitudinal count," to be even. Because of the non-return-to-zero type of recording used, the check character can be generated simply by resetting the current in the writing heads to the reference polarity; this has the added advantage of allowing the interrecord gaps always to be saturated to a known polarity in all tracks, so that a record may easily be rewritten if an error occurs in recording.

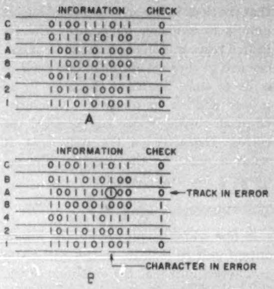


Fig. 16. The method of detecting errors on magnetic tape. In (A) the recording is correct, as indicated by the even count of "1s" in each character and each track. In (B) the "A" bit of the third character from the end of the record is incorrect

The manner in which the longitudinal check is used may be seen from Fig. 16 (A), a properly recorded tape record; and 16(B), a record in which a "1" appears erroneously in place of a zero in the encircled position. This could result from an error in reading, or from damage to the tape after it is recorded. In either case, the operator will receive a signal indicating the track in error, and by a simple operation from the control console, will be able to determine the character in error by the normal redundancy check, and subsequently correct the error within the memory by inverting the bit which is incorrect. Incidentally, the use of both transverse and longitudinal checks on the tape allows detection of single or double errors as well as correction errors.

PHILOSOPHY OF CHECKING

In case of the detection of an error of any kind, the appropriate signal line is energized and becomes available for in-

Vol.
Pt. 1
1955

terrogation by the TRANSFER ON SIGNAL instruction. This conditional transfer allows corrective action to be taken automatically, so that the 702 may proceed with its job.

If an error occurs in reading a record from tape, for example, there is a good possibility that the record would be read correctly on a second try. It has therefore been found better to let the reading process be repeated by the program, upon detection of an error, than to introduce possible further errors by manual intervention. This reasoning holds for the other classes of errors as well.

Conclusion

The ability of the 702 to handle data in words of variable lengths introduces the requirement of special control circuits in the form of execution timers. The arithmetic and logical unit, in fact, is made up very largely of memory circuits and the kind of logical circuits required to control the execution of serial operations upon words of different lengths. The amount of equipment required for the actual arithmetic functions, such as addition, on the other hand, is relatively small because of the serial nature of the machine.

The very large arithmetic accumulator makes possible several manipulative operations, including the transfer of large blocks of data within the arithmetic and logical unit, which are not feasible in conventional computers.

Also basic to the design of the 702 is the expandable input-output system, with each new unit introducing its own power supply. As improved versions of individual input and output devices are made available, they can be connected with no modification of the existing equipment.

Reliability and ease of servicing were given prime consideration in the design of the equipment, and voltages for "marginal" testing were selected early in the circuit design phase. Semiautomatic selection of circuit sections and variation of test voltages make possible a complete check of the machine in a very short time.

Appendix I. Character Code Chart

The 7-bit code for the 702, shown in Table III, is grouped in four columns in order to show correspondence with the standard IBM punched-card code. The four columns from left to right, with zoning (*BA* bit combination) 11, 10, 01, and 00, correspond to the zoning 12, 0, and "blank" respectively, in punched cards.

	C BA 8421		C BA 8421		C BA 8421		C BA 8421
&	0 11 0000	—	1 10 0000	BLANK	1 01 0000	ACC. MARK	0 00 0000
A	1 11 0001	J	0 10 0001	/.	0 01 0001	1	1 00 0001
B	1 11 0010	K	0 10 0010	S	0 01 0010	2	1 00 0010
C	0 11 0011	L	1 10 0011	T	1 01 0011	3	0 00 0011
D	1 11 0100	M	0 10 0100	U	0 01 0100	4	1 00 0100
E	0 11 0101	N	1 10 0101	V	1 01 0101	5	0 00 0101
F	0 11 0110	O	1 10 0110	W	1 01 0110	6	0 00 0110
G	1 11 0111	P	0 10 0111	X	0 01 0111	7	1 00 0111
H	1 11 1000	Q	0 10 1000	Y	0 01 1000	8	1 00 1000
I	0 11 1001	R	1 10 1001	Z	1 01 1001	9	0 00 1001
PLUS ZERO	0 11 1010	MINUS ZERO	1 10 1010	REC. MARK	1 01 1010	0	0 00 1010
.	1 11 1011	$	0 10 1011	,	0 01 1011	#	1 00 1011
⌑	0 11 1100	*	1 10 1100	%	1 01 1100	@	0 00 1100
						TAPE MARK	0 00 1111

Table III. Character code chart for the IBM type 702

The code of Table III departs in only two places from a direct logical translation from the 12-bit punched-card code to the 7-bit code. On punched cards, the character in the upper right-hand corner of Table III would be the blank; however, since the blank must be representable on magnetic tape, it was moved one place to the left, and the "all-zeros" combination was reserved for a character which does not appear on tape.

It was desirable to place the zero in the right-hand column along with the other numerics, partly in order to allow the "plus" and "minus" signed zeros to be in the same row with numeric zero.

The tape mark is a special mark recorded at the end of recording on a tape reel to indicate the end of the file when reading. It turns on the input-output indicator for the selected tape unit when it is encountered during the READ instruction.

Appendix II. List of Operations of the 702

The letters or numerals in parentheses signify the operation part.

ADD (G). Add specified numeric memory field to accumulator field algebraically. Length of result in accumulator equals length of longer of two fields added, unless overflow increases length by one digit. Time to read and execute instruction: 0.368 millisecond for 10-digit fields.

SUB (P). Subtract specified numeric memory field from accumulator field algebraically. Same length rule and time as ADD.

R ADD (H). In effect, reset accumulator field to zero and add specified numeric memory field, transfering numeric fields to accumulator and setting accumulator sign according to sign of memory field.

R SUB (Q). Reset accumulator field and subtract specified numeric memory field algebraically, setting accumulator sign appropriately.

ADD MEM (6). Add accumulator field to specified memory field algebraically. If memory field starts with a signed digit in lowest-order position, memory field is considered numeric and addition continues until another signed digit is encountered. If memory field starts with an unsigned digit, addition includes zone bits and continues through length of accumulator field. Latter case prevails in adding to instructions.

MPY (V). Multiply specified numeric memory field by accumulator field, leaving product in accumulator. Product length equals sum of multiplier and multiplicand lengths. Time to read and execute instruction: 3.36 milliseconds for 10-digit factors.

DIV (W). Divide accumulator field by specified numeric field from memory, leaving quotient field in accumulator. Length of quotient field equals length of dividend minus length of divisor field. Time: 21.5 millisecond for 10-digit divisor and quotient fields.

SHOR (C). Shorten accumulator field by deleting from its right end specified number of characters, and moving starting point counter to new lowest-order position.

LENG (D). Lengthen accumulator field by adjoining specified number of zeros to its right end and moving starting point counter to new right end of field.

ROUND (E). Round off accumulator field by deleting specified number of digits, but add 5 to highest-order digit deleted, propagating any carries and allowing overflow, if any, to add one position to resulting field size.

SET L (B). Set left end of accumulator field by placing accumulator mark so as to effect field length specified in instruction. Any positions added to accumulator field are filled with zeros. Time: same as

ADD for equal resulting accumulator field length.

NORM TR (X). Normalize and transfer. If left-hand digit in accumulator is zero, it is deleted and a transfer of control to the specified address is effected. Used in normalizing factors in floating decimal operation.

STORE (F). Store numeric accumulator field in specified memory location, with plus or minus zoning over units digit to correspond with sign of accumulator. When accumulator mark is reached, corresponding memory character (one position to the left of field being stored) receives plus zoning if it is an unsigned numeric. Otherwise this character is left unaltered.

ST PR (5). Store numeric accumulator field in specified memory location for printing. Where commas or decimal points are encountered in memory, they are left unchanged, digits from accumulator being stored in free spaces in preset memory pattern. All zeros and commas to left of highest-order nonzero, nonperiod character are deleted.

SIGN (T). Set accumulator sign according to zoning of character in specified memory position, leaving character with numerical zoning. Accumulator sign is left negative if zoning is minus, and positive otherwise. Accumulator is left with 1-character field; "&" if plus and "−" if minus.

COMP (4). Compare accumulator field with specified memory field. Fields may be numerical or alphabetic. Any zoning over lowest-order character in memory is considered part of character to be compared, not as sign. Comparison is absolute, and accumulator sign is disregarded. High or equal trigger may be set according to result of comparison.

LOAD (8). Enter memory information into accumulator, starting at specified memory address and stopping when accumulator mark is read. Accumulator field length remains unchanged. Memory field is considered alphabetic or mixed, sign being disregarded. Accumulator sign is set plus.

UNLOAD (7). Reverse of LOAD.

TR (1). Unconditional transfer of control to instruction at specified address.

TR HI (K). Transfer if accumulator field was higher than memory field in most recent COMP instruction.

TR EQ (L). Transfer if accumulator and memory fields were equal in most recent COMP instruction.

TR PLS (M). Transfer if accumulator sign is plus.

TR ZRO (N). Transfer if accumulator sign is zero.

TR SIG (O). Transfer if selected signal line is on.

SEL (2). Select specified input or output device for reading or writing; or select specified input-output indicator, check indicator or program alteration switch for interrogation.

READ (Y). Read from last selected input device into memory, starting at position specified and proceeding toward right. Store a record mark in memory after last character of record has been read.

WRITE (R). Write from memory to output device last selected, starting at position specified and proceeding toward right, until record mark is encountered.

R CHK (I). Read and check. Read from last selected input device and compare with memory, starting at position specified and proceeding toward right. If comparison shows unequal records, turn on the read-write check indicator.

WR ER (Z). Write and erase. Same as WRITE, but erase memory as writing proceeds, replacing memory characters with blanks.

CTRL (3). Control last selected input or output device by performing the function specified in the address part of the CTRL instruction.

STOP (J). Stop machine at program step containing STOP instruction.

NO OP (A). No operation. Machine proceeds to next instruction.

References

1. The IBM Type 702, an Electronic Data Processing Machine for Business, C. J. Bashe, W. Buchholz, N. Rochester. *Journal*, Association for Computing Machinery, Pittsburgh, Pa., vol. 1, no. 4, Oct. 1954, pp. 149–69.

2. Improved Cathode-Ray Tube for Application in Williams Memory System, W. E. Mutter. *Electrical Engineering*, vol. 71, Apr. 1952, pp. 352–56.

3. Engineering Experience in the Design and Operation of a Large Scale Electrostatic Memory, J. C. Logue, A. E. Brennemann, A. C. Koelsch. *Convention Record*, Institute of Radio Engineers, New York, N. Y., pt. 7, 1953, pp. 21–29.

4. Review of Input and Output Equipment Used in Computing Systems. *AIEE Special Publication S-53*, "Magnetic Tape Recording Techniques and Performance," H. William Nordyke, Jr., Mar. 1953, pp. 90–95.

Developments in Closed-Circuit Television

MAX H. KRAUS
ASSOCIATE MEMBER AIEE

THE development of closed-circuit television (TV) has been marked by many fascinating applications but, until recently, economic factors ruled out its mass usage in many fields. In industry, education, business, and the military, television has been a fascinating medium to explore, test, and evaluate, but only infrequently could its proponents justify its introduction on a permanent large-scale basis. Recently, however, advances in two fields have made this possible.

The first development was the perfection of the Vidicon pickup tube, and the subsequent introduction of Vidicon cameras selling for a fraction of the cost of studio-type image orthicon units. Prices have been reduced to the point so that it is possible to make an installation for between $1,000.00 to $5,000.00, with features and performance to meet almost any task. With this has come a host of familiar "single-camera" applications. Industry has quickly accepted TV as a means for watching loading platforms, smoke stacks, rolling mills, and scores of other situations where it is essential for a single operator to know the action taking place at a single remote point. Educators and business men have found similar needs. Important though these applications are, the limits placed on the number of cameras and pickup-reception points which can be economically and effectively covered by an elementary installation for a time, put a virtual ceiling on the extent of TV applications.

The development of the flexible TV distribution system has eliminated this barrier. It is now possible not only to transmit simultaneously a number of signals inexpensively but to install cameras either permanently or insert portable cameras at many points in a system, providing ultimate flexibility while maintaining complete control over each channel.

Before investigating these new developments, it is well to consider some of the costs heretofore associated with the larger closed-circuit installations. Assuming a need for a high-quality, interlaced picture, the camera with a viewfinder, lens, and tripod cost between $3,000 and $5,000. The TV receivers added between $125 and $250 per viewing location while installation costs in a 10-set single-channel radio-frequency or video-distribution system, within a medium-sized building, varied from $200 to $1,000. With the addition of $500 for mainte-

---

Paper 55-687, recommended by the AIEE Television and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted June 6, 1955; made available for printing July 28, 1955.

Max H. Kraus is with the Jerrold Electronics Corporation, Philadelphia, Pa.

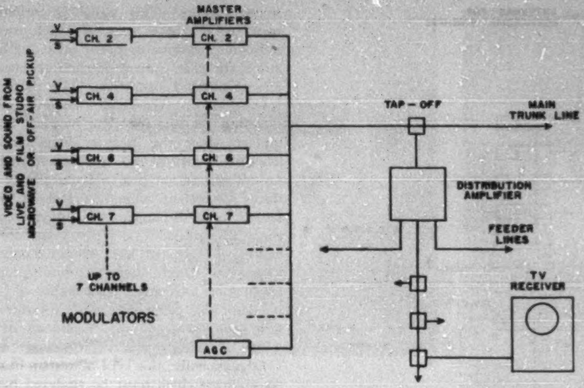

MASTER AMPLIFIERS
VIDEO AND SOUND FROM LIVE AND FILM STUDIO MICROWAVE OR OFF-AIR PICKUP
V
S
CH. 2
CH. 4
CH. 6
CH. 7
UP TO 7 CHANNELS
MODULATORS
AGC
TAP — OFF
MAIN TRUNK LINE
DISTRIBUTION AMPLIFIER
FEEDER LINES
TV RECEIVER

Fig. 1. Basic TV distribution system

nance and parts for the first year, the total cost quickly reached between $5,000 and $9,000. Furthermore, if the system was used to watch animated scenes such as are found in education, the cost of one or more cameramen or operators was added plus the studio or other central control requirements. Finally, moves or changes presented phenomenal problems. The setup and rebalancing time alone made it feasible to move the camera or the receivers only for events of the first magnitude.

Multichannel TV Distribution Systems

It was against this background that the flexible, multichannel TV distribution system was introduced. As shown in Fig. 1, a basic system consists of modulators or, as they are sometimes called, miniature transmitters, a master amplifier, coaxial cable, distribution equipment, and receiver-connection accessories. The TV installation at Fort Monmouth, N. J., utilizes a large system of this type to illustrate both a military and an educational application. As shown in Fig. 2, both live and film programs are originated in the studios and distributed throughout the training areas of the post by a 3-channel system operating on channels 2, 4, and 6. Over 100 classrooms, various theatres, and the field house are served so that students receive both training in their studies and general orientation-information-type programs. In addition, programs picked up from commercial stations are presented as well as "remotes" microwaved from other parts of the post. Evidence points to the fact that thousands of lost training hours have been eliminated since it is no longer necessary to move troops from one area to another to receive instruction. Furthermore, films can be shown without darkening the room—students stay more alert and can take notes—and the teaching continuity is not broken as is the case when the instructor takes time to set up and operate the movie projector or moves the class from one room to another for a viewing. Films last longer and many duplicate films and projectors can be eliminated. Thus, the TV distribution system with its multichannel capabilities and classroom outlets has become an invaluable training medium.

In industry or business, this same type of system can be used in many ways, including providing scheduling and status information from a central point to many subordinate points. In other cases, the closed-circuit function can be combined with master TV antenna system needs to provide both air channels and closed-circuit channels over the same cable. Security needs can be met too. Techniques which will be discussed in the following make it possible to restrict certain channels to be viewed at specific

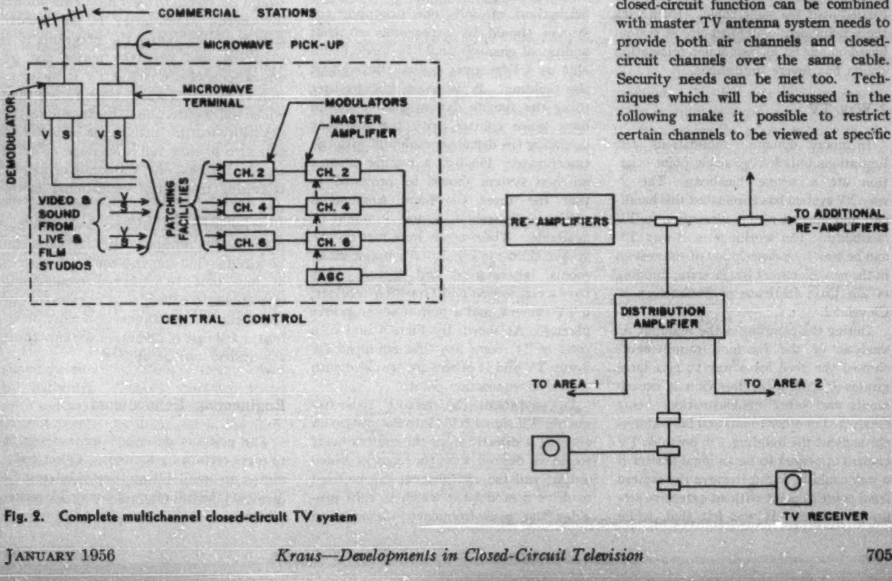

COMMERCIAL STATIONS
MICROWAVE PICK-UP
DEMODULATOR
MICROWAVE TERMINAL
V
S
V
S
MODULATORS
MASTER AMPLIFIER
VIDEO & SOUND FROM LIVE & FILM STUDIOS
PATCHING FACILITIES
CH. 2
CH. 4
CH. 6
CH. 2
CH. 4
CH. 6
AGC
CENTRAL CONTROL
RE-AMPLIFIERS
TO ADDITIONAL RE-AMPLIFIERS
DISTRIBUTION AMPLIFIER
TO AREA 1
TO AREA 2
TV RECEIVER

Fig. 2. Complete multichannel closed-circuit TV system

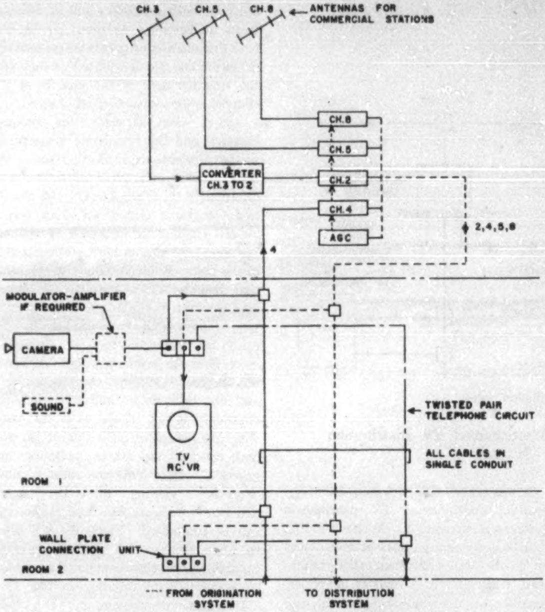

CH.3
CH.5
CH.8
ANTENNAS FOR COMMERCIAL STATIONS
CH.8
CH.5
CONVERTER CH.3 TO 2
CH.2
CH.4
AGC
2,4,5,8
4
MODULATOR-AMPLIFIER IF REQUIRED
CAMERA
SOUND
TV RC'VR
TWISTED PAIR TELEPHONE CIRCUIT
ALL CABLES IN SINGLE CONDUIT
ROOM 1
WALL PLATE CONNECTION UNIT
ROOM 2
FROM ORIGINATION SYSTEM
TO DISTRIBUTION SYSTEM

Fig. 3. Basic 2-way TV system, unattended central facilities

receivers, even though all channels are carried over the same cable. Even with these refinements, a closed-circuit system based on one or, at best, a few origination points has severe limitations.

2-Way TV

In many dynamic applications, the limitations imposed by single-point origition are a severe handicap. The 2-way TV system has eliminated this handicap and opened a new dimension in TV flexibility. The workings of 2-way TV can be seen by a description of the system in the new Electrical Engineering Building of the Case Institute of Technology in Cleveland.

During the planning of the building, an analysis of the teaching requirements showed the need for a way to give large groups of students a close view of experiments and other demonstrations being conducted in other rooms and laboratories throughout the building. A portable TV camera appeared to be an ideal answer if a way could be found to move the camera from point to point without extensive set-up procedures. It was felt that, while initial provision should be made for the origination of only one program, the system should be expandable so that additional channels could be made available at TV receiver outlets throughout the building. In addition, the lecturer using the remote demonstration should have voice contact with the assistant operating the demonstration and with the cameraman. Finally, a regular master antenna system should be provided so that the three Cleveland commercial stations, channels 3, 5, and 8, would be available. These needs were met by the system shown in Fig. 3. All major classrooms, laboratories, and lecture halls have a connection point for a TV receiver, a TV camera, and a pair of sound power phones. As shown by Figs. 4 and 5, a total of 34 rooms are thus equipped for 2-way TV and 11 others are provided with a receiver-connection point.

In operation, the remote radio-frequency TV signal is fed into the system on channel 4 directly from the camera or, if sound is desired too, the camera video output and the audio signal can be used to drive a modulator which in turn provides the radio-frequency picture and sound carriers. This signal is carried back to the master amplifier which uses individual channel amplifier strips with automatic gain control on each channel for unattended operation. This latter feature allows the same camera (or modulator) output level to be used no matter where the unit is plugged in. The availability of this universal setting for all origination points drastically reduces the set-up time. In fact, it is merely a case of plugging the camera into the wall terminal to have it ready for action. The system has been in use only a short time, but already a number of applications have arisen, such as the building dedication ceremonies, when the overflow crowd watched the ceremonies by TV in adjoining rooms.

Educationally, this type of system has possibilities which must be analyzed by the educators. However, several points are clear. Regarding cost, it eliminates the studio requirements entirely and, instead of requiring highly trained personnel, it can be operated by a laboratory assistant or student. Tests have shown that, from the production point of view, the program need not measure up to commercial standards.[1] Functionally, it can be used either as a teaching aid or a teaching medium. As an aid, it can bring large objects into the class-room or magnify objects too small to be seen simultaneously by all students. As a teaching medium, it enables the professor to remain in his natural habitat, and would thus seem to meet most educational TV needs. The latter is particularly true if in evaluating the programs it is kept in mind that the major goal of educational closed-circuit TV is the creation of programs which teach rather than which sell or entertain. Furthermore, the flexibility feature enables the educator to educator to take full advantage of all of the possibilities of TV instead of merely televising traditional educational methods.

In industry and business, 2-way TV has many as yet untapped uses. Plants can be wired so that cameras can be plugged in at any time to observed troublesome operations, to collect data, or even to tour the facilities. Wherever low cost, flexibility, and ease of operation are important the system can be applied.

Engineering Refinements

The systems described were designed to meet certain basic needs. Other techniques are available and are of interest to illustrate the full range of 2-way TV possibilities.

Vol.
Pt. 1
195

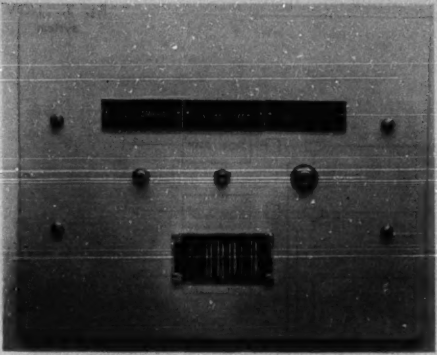


Fig. 4 (above). Wall connection point for TV camera, receiver, and telephone

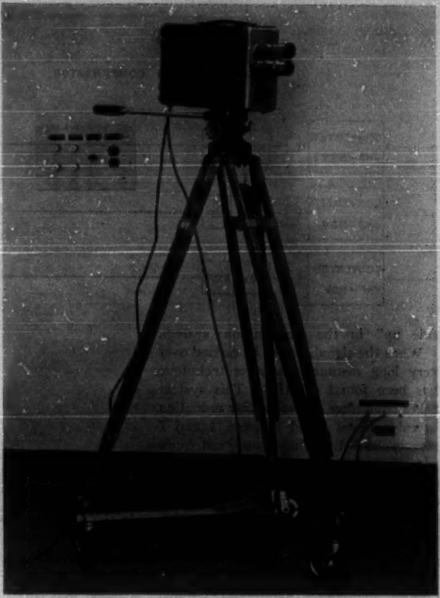


Fig. 5 (right). Vidicon camera connected to system

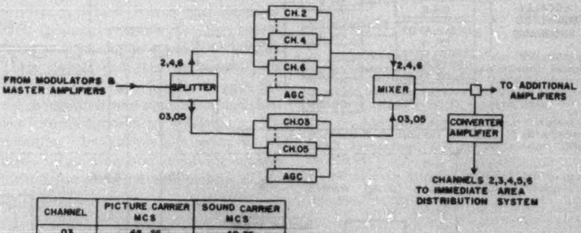
ol. 74
. 1
955

The Fort Monmouth system covers an entire military post and uses approximately 5 miles of cable to reach over 100 outlets. The Case Institute system, although planned for campuswide expansion, covers only a single building with only hundreds of feet of cable to reach its 45 receiver outlets. As might be imagined, the wide range of TV distribution system sizes has brought about several different system approaches. The entire subject of master antenna systems and community antenna systems has been given detailed treatment,[2,3] but the problem of system size is sufficiently important to deserve mention.

In a small system, although the problems are relatively simple, the need for automatic-gain control equipment, proper signal levels, etc., is apparent. However, in a large system, additional problems are encountered in designing equipment which will provide adequate bandwidth and flat frequency response through many cascaded amplifiers. Furthermore, the need to minimize equipment makes it desirable to stay below 100 megacycles since, because of cable losses, if high TV channels (7 to 13) are used, the equipment requirements increase by roughly a factor of 4 for amplifiers of the same gain and output. These additional factors have created the need for two types of systems: the signal distribution system and the signal transport system.

The Case Institute system utilizes only signal distribution, which is accomplished by individual channel amplifier strips with automatic-gain control located at a central point. As discussed in the foregoing, both low channels (2, 4, and 5) and a high channel (8) are used. Compensation for the difference in attenuation between channel 2 and channel 8 is accomplished by setting the output level of the channel-8 strip higher than that of the channel-2 strip, and by using capacitive tap-off units which have a lower attenuation in the high channel than in the low.

A 5-channel transport system is shown in Fig. 6. This system uses channels 2, 4, and 6, to which have been added channels designated as 03 and 05 occupying bands of 40.5 to 46.5 megacycles and 24.5 to 30.5 megacycles respectively. The relation between the two subchannels is determined by the desire to convert these to channels 3 and 5 for eventual distribution to the TV receivers. The operation of this system imposes no unusual requirements. Only one departure from normal techniques is required. The conversion from channels 03 and 05 to channels 3 and 5 by a single oscillator operating above channel 5 inverts the picture and sound carrier positions. It is therefore necessary, when originating a program at the 03 or 05 frequencies, to originate it with the picture and sound portions inverted so that they come out with their "right



CHANNEL	PICTURE CARRIER MCS	SOUND CARRIER MCS
03	45.25	40.75
05	29.25	24.75
2	55.25	59.75
4	67.25	71.75
6	83.25	87.75

Fig. 6. Five-channel transport system

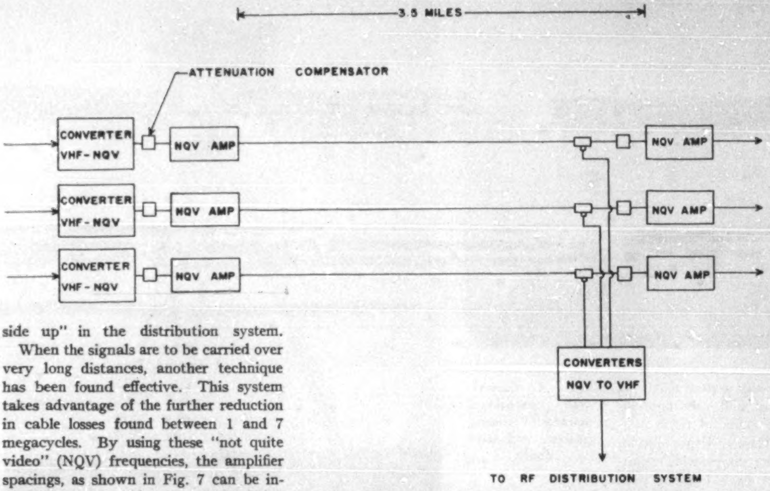

3.5 MILES
ATTENUATION COMPENSATOR
CONVERTER VHF-NQV
NQV AMP
CONVERTER VHF-NQV
NQV AMP
CONVERTER VHF-NQV
NQV AMP
NQV AMP
NQV AMP
NQV AMP
CONVERTERS NQV TO VHF
TO RF DISTRIBUTION SYSTEM

Fig. 7. NQV system

side up" in the distribution system.

When the signals are to be carried over very long distances, another technique has been found effective. This system takes advantage of the further reduction in cable losses found between 1 and 7 megacycles. By using these "not quite video" (NQV) frequencies, the amplifier spacings, as shown in Fig. 7 can be increased to over $3^1/_2$ miles of *RG-11/U*-type cables. Since all channels are between 1 and 7 megacycles, a separate cable must be used for each. However, installation labor is not appreciably increased by adding additional parallel cables, so that with the reduction in the amount of electronic equipment needed the system has many attractive, low-cost features.

In all types of signal transport systems, the type of coaxial cable is of extreme importance. Recently developed cables having a more uniform frequency response and lower temperature coefficient of attenuation have materially assisted in the successful operation of long-signal transport systems. Furthermore, these new cables, known as *JEL*-type, utilize a high molecular-weight, noncontaminating, polyethylene jacket in place of the traditional vinyl jacket used on most *JAN*-type cables. This new jacket materially increases cable life and, therefore, system life.

The transport system technique has further reduced the cost of closed-circuit systems. For example, if distribution of only two channels is required, but over a large area, channels 03 and 05 can be used. This minimizes equipment and maintenance since it is merely necessary to provide inexpensive converters at the TV receivers to convert channels 03 and 05 to regular very-high-frequency channels.

Channels 03 and 05 can be used in another way. The Case Institute System is a double-cable system—one for origination, one for distribution. If only a few channels are required, it is possible to send the remote programs from the camera to the central amplifier on channels 03 and/or 05, while distributing these or other programs on channels 2, 4, and 6—all over the same cable. As shown in Fig. 8, this technique has the

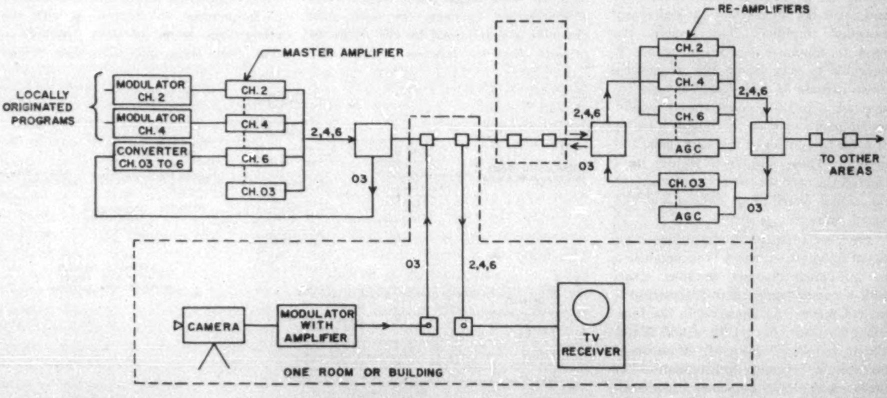

MASTER AMPLIFIER
RE-AMPLIFIERS
LOCALLY ORIGINATED PROGRAMS
MODULATOR CH. 2
MODULATOR CH. 4
CONVERTER CH. 03 TO 6
CH. 2
CH. 4
CH. 6
CH. 03
2,4,6
03
2,4,6
03
CH. 2
CH. 4
CH. 6
AGC
CH. 03
AGC
2,4,6
03
TO OTHER AREAS
03
2,4,6
CAMERA
MODULATOR WITH AMPLIFIER
TV RECEIVER
ONE ROOM OR BUILDING

Fig. 8. Single-cable 2-way TV system

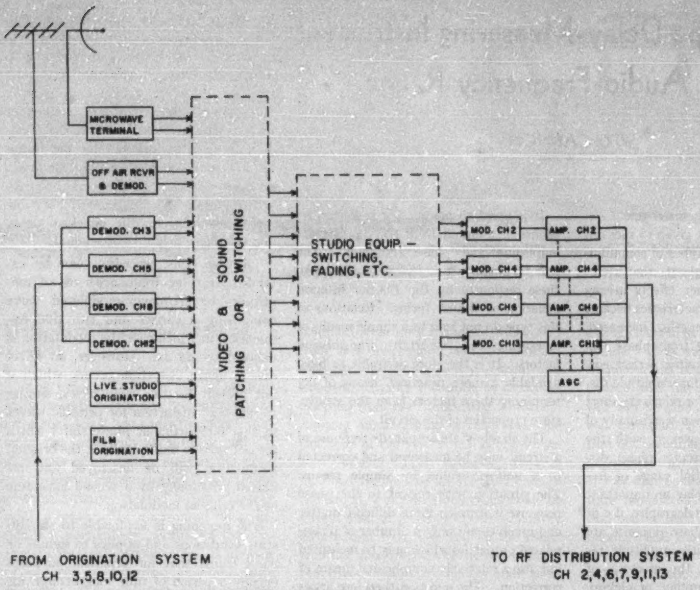

MICROWAVE TERMINAL
OFF AIR RCVR & DEMOD.
DEMOD. CH3
DEMOD. CH5
DEMOD. CH8
DEMOD. CH12
LIVE STUDIO ORIGINATION
FILM ORIGINATION
VIDEO & SOUND PATCHING OR SWITCHING
STUDIO EQUIP. — SWITCHING, FADING, ETC.
MOD. CH2
MOD. CH4
MOD. CH6
MOD. CH13
AMP. CH2
AMP. CH4
AMP. CH6
AMP. CH13
AGC
FROM ORIGINATION SYSTEM CH 3,5,8,10,12
TO RF DISTRIBUTION SYSTEM CH 2,4,6,7,9,11,13

Fig. 9. Central control facilities, attended operation

advantage of eliminating the second cable and is particularly valuable in providing 2-way TV coverage over a large area with a minimum of cable.

The sub channels have even another use, namely, program zoning. For example, it may be desirable to distribute three programs to all TV receivers in an area, and to have two additional programs available to only certain ones. This can be accomplished by distributing channels 03, 05, 2, 4, and 6. All receivers then can view channels 2, 4, and 6, but only those equipped with a special converter can receive the restricted 03 and 05 programs.

When desired, another form of control can be instituted—program selection. It is possible to introduce timed, automatically sequenced switching and control equipment at the head end of the system to select only certain programs for distribution. This feature is particularly adapted to hospitals where a close watch is placed on the program content reaching disturbed patients and in schools where indiscriminate use of television might be objectionable.

No mention has been made of studio control equipment because in most cases none is required. However, some applications require programs originated both by 2-way TV and by studio techniques, either film or live. This is particularly true where central movie telecasting facilities are provided. In this case, the automatic design, whereby a program originated at a remote point and is automatically distributed to the entire system, is eliminated. In its place, demodulator units are provided, as shown in Fig. 9, which reduce the incoming radio-frequency signals to the lowest common denominator of video and sound. These signals can then be switched or combined with other signals, through the use of regular studio equipment, and then fed into modulators for subsequent distribution. In this way a program can be made up of sequences involving remote pickup, live pickup, studio pickup, slides, and film.

The Future

The future of closed-circuit TV is bright. Systems can be devised to meet almost any needs, especially because closed-circuit systems can be engineered unfettered by restrictions on spectrum space and compatibility. For example, special color systems and systems giving pictures with increased resolution are only two possibilities. Today one fact is apparent: the engineering techniques have been developed and the time is ripe for the applications to find the market.

References

1. Television in Army Training: Evaluation of Television in Army Basic Training, Joseph H. Kanner, Richard P. Runyon, Otello Desiderato. *HumRRO Technical Report 14*, George Washington University, Washington, D. C., Nov. 1954.

2. Line Amplifiers for Community Television Systems, K. A. Simons, Don Kirk, H. J. Arbeiter. *AIEE Transactions*, vol. 73, pt. I, Apr. 1954, pp. 214–20.

3. TV Master Antenna Systems, Ira Kamen, Richard Dorph. John F. Rider Publications, New York, N. Y., 1951.

An Envelope Delay-Measuring Instrument in the Audio-Frequency Range

W. D. CANNON
FELLOW AIEE

IT HAS long been recognized that careful attention must be given to the amplitude response of all types of communication systems throughout the useful frequency range. Other effects arising from system phase characteristics became evident as circuit lengths increased. These effects, resulting from phase distortion, persisted even after perfect correction had been made for amplitude distortion. In telephony for relatively short distances, phase distortion was usually of small importance. However, with the increased length of circuits phase distortion over the essential range of frequencies was found to play an important role. In landline code telegraphy, d-c or carrier, the effects of phase response are negligible compared with amplitude response. In any event, the processes of signal regeneration, sampling, or pickup, and code translation provide a "margin" within which distortion effects are absent and a perfect printed reproduction of the telegraph message is obtained.

With the advent of facsimile, telephotography, and television, it was soon apparent that these systems, lacking the compensatory factors present in code telegraphy and telephony, require careful consideration of phase relations as well as amplitude response. Thus, in facsimile transmission the signals are largely transient in nature and for faithful reproduction in the received copy, the wave shape of the transients must be preserved or restored to the original form prior to recording.

It is usual to predict the transient response of a circuit from the steady-state amplitude and phase responses. The transient response is uniquely related to these responses by the Fourier integral equation or similar forms. Relations of this type do not lead to a simple means of measuring the important transmission factors. It is therefore desirable to have available a more practical means of determining these factors from the steady-state properties of the circuit.

The steady-state amplitude response of a circuit may be measured and corrected to a uniform value by simple means. The situation with regard to the phase response is a much more difficult matter and involves not only a number of phase-related properties which may be measured but also a relatively complicated means of correction. The question therefore arises as to what phase-related parameters or properties to measure, particularly for corrective purposes. This paper considers this question and describes a new instrument for measuring envelope delay in the audio-frequency range.

Phase-Related Parameters

In the transmission of periodic waves through a medium, the amplitude and phase responses are usually considered in terms of the steady-state parameters α and β so that the propagation constant is $\alpha+j\beta$. Since periodic pulses may be represented by a constant along with a number of harmonically related sine and cosine waves having different amplitudes, the shape of the pulses will be preserved in transmission if all the components are transmitted with the same velocity at a uniform attentuation α of the circuit. For periodic waves the condition for transmission without phase distortion is that $\beta=\omega\tau+m\pi$ where τ is the time delay interval in transmission and m is an integer. If m is not an integer, a form of distortion is present known as linear-phase intercept distortion.

If, instead of periodic waves, two sinusoidal waves of slightly different frequencies are considered as being propagated through a medium, it can be shown that the conditions for distortionless transmission are somewhat different. These two waves will travel through a medium having phase distortion at different velocities and will form crests and troughs which will travel at a slightly different velocity from the component waves. It can be shown that the crests thus formed are delayed by the factor $d\beta/d\omega$ commonly referred to as the group or envelope delay, whereas the component waves are delayed by β/ω. For a distortionless medium, the envelope delay and the phase delay become equal.

Likewise, this procedure can be extended to three frequencies of an amplitude- or frequency-modulated wave comprising a carrier and two side frequencies, in which case the modulation is delayed by $d\beta/d\omega$. However, an effect due to phase distortion occurs, which is not present for two frequencies, similar to intercept distortion for periodic waves in that for particular phase shifts which are not zero or multiples of π, the original modulation will be destroyed and replaced principally by a second harmonic of the original modulation.

The foregoing is applicable to steady-state conditions and applies to signals or wave trains that for practical purposes occupy a period of time long enough for all transient effects to have disappeared. A study of finite wave trains rather than infinite wave trains runs into considerable difficulty even for the most practical simplifying assumptions. The transmission of a few types of finite wave trains has been studied by Carson[1] and Meade.[2] Here, not only the first derivative $d\beta/d\omega$ but also derivatives of higher order are of importance.

Hence, with respect to delay distortion in transmission circuits, the most important factor to be considered in general is envelope delay $d\beta/d\omega$, and it is not necessary to consider phase shift directly as such. Further, in a distortionless system, the phase delay β/ω and the envelope delay $d\beta/d\omega$ are equal and constant throughout the useful frequency range and the second derivative $d^2\beta/d\omega^2$ is zero. Since the ideal system is generally not available in practice, the phase delay and the envelope delay will not be equal and the second derivative will not be zero. Fortunately, however, when satisfactory correction has been made for envelope delay distortion, usually the phase delay also is within tolerable limits unless the phase of the individual carrier cycles under the modulation is to be used. In this case β/ω should conform to some prescribed tolerance depending upon the actual application, particularly in regard to linear-phase intercept distortion.

Vol
Pt.
195

Paper 55-734, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3-7, 1955. Manuscript submitted June 7, 1955; made available for printing August 8, 1955.

W. D. CANNON is with the Western Union Telegraph Company, New York, N. Y.

The development of the instrument described in this paper was one phase of a comprehensive investigation of facsimile signal transmission conducted by the Western Union development and research organization under U. S. Army Signal Corps sponsorship. The author gratefully acknowledges the guidance and assistance of many people concerned with the measurement of envelope delay and measurement, and correction of delay distortion. Special thanks are given to Henry F. Burkhard and Chester S. Friedman of the U. S. Army Signal Corps, and to F. B. Bramhall, transmission research engineer, D. J. Bertuccio, Harold Glick, E. S. Grimes, and T. Rystedt of The Western Union Telegraph Company.

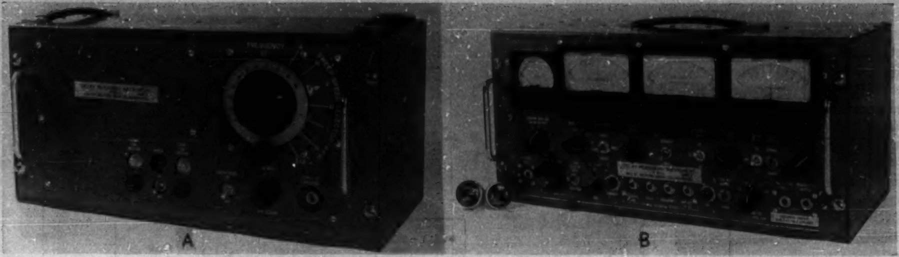


Fig. 1. A—Unit 1: regulated power supply and sweep oscillator. B—Unit 2: transmitter and receiver

Methods of Measuring Envelope Delay

It is evident that the use of steady-state conditions for the measurement of envelope delay in audio-frequency circuits has practical advantages both in simplicity of techniques and instrumentation, whereas measurements as a function of some quantity related to the transient state appear to be confronted with difficulties, particularly because of the connecting integral relationships. A number of methods of measuring envelope delay in terms of steady-state conditions are available, some of which give the result indirectly and others the result in the form of a characteristic curve. Indirect methods of measurement usually do not require elaborate equipment, but do require that the results be derived from the measurements, resulting in a somewhat complicated procedure. One procedure which may be used is the direct measurement of the phase shift itself, which may be obtained by a number of methods. The resulting curve may then be differentiated with respect to frequency to yield the envelope delay. Another method of measurement is by means of a single frequency which is varied in small steps or increments such that the incremental change $\Delta\omega$ can be considered essentially as $d\omega$. The corresponding $\Delta\beta$ is obtained by measuring the change in phase angle on a phase meter or similar device. The delay curves then are $\Delta\beta/\Delta\omega$ which are essentially equivalent to $d\beta/d\omega$ as obtained by a point-by-point method.

The methods just discussed are in general not very practicable for incorporation in an instrument for continuous measurements of delay particularly when it is desired to display the result as a plot of delay versus frequency. None of these methods are practicable for measurements on transmission systems in which voice frequency circuits are derived by a modulation process requiring translation from one frequency band to another, since the received frequency in such voice frequency circuits may differ by a number of cycles from the transmitted frequency. Another disadvantage arises when the two ends of the system are isolated from each other. This situation requires matched frequencies at the two ends of the circuit, which in many cases would be impractical, or another circuit to convey information from one end to the other. Another disadvantage of these methods is the time-consuming point-by-point process usually required.

A more practicable method, by which the result is obtained directly, is available by the use of two sinusoidal waves of slightly different frequency, which, as discussed previously, yields envelope delay. The result is obtained by measuring the time relationship of the crests of the envelope thus formed with respect to a reference, as the two waves are swept through the frequency range. The two waves, having a constant and small frequency difference, can be obtained by a modulation process. The delay information is then obtained by measuring the time relationship of the envelope as recovered by a demodulation process at the receiving end with respect to the original modulating frequency at the transmitting end.

This method is practicable when both ends of the circuit are available at one point or when a separate circuit is available between two separate points, but is not very practicable for the usual voice frequency circuit when the two ends of the circuit are in isolated localities. A more practicable method is available when three frequencies instead of two are transmitted; the third frequency, a carrier with respect to the other two, may be used to supply information to provide a sweep for displaying delay at the receiving end against the transmitted frequencies as received at the receiving end of the circuit. It is evident that these three frequencies as obtained by a modulation process will have phase coherence and will add to give crests which in general will travel at a different velocity from that of the component waves due to delay distortion, thus yielding envelope delay. It can also be shown that either amplitude modulation or frequency modulation can be used to derive the three frequencies and that this will yield the desired result. However, since frequency modulation appears to be more difficult for instrumentation at the low-carrier and modulation frequencies involved and since signals to be transmitted such as facsimile signals are generally amplitude-modulated, the amplitude modulation method of deriving the three frequencies appears to be more practicable for use in a portable instrument.

The use of the amplitude modulation method of deriving a carrier and two side-band frequencies avoids the shortcomings described in the foregoing and permits the two ends of the circuit to be in isolated localities. The carrier at the transmitting end is modulated with a low-frequency test signal, such as 25 cycles, derived from a stable standard source. It is necessary only that this modulated wave be transmitted over the circuit, and from it the desired information on delay and amplitude characteristics can be obtained. At the receiving end, the test signal is recovered from the modulated wave and is compared with a local standard on a relative time basis. If the relative time indicating device has a linear response to this time relationship, then the standard sources at each end of the circuit need not be synchronized or phased with each other but may be allowed to drift slowly with respect to each other without limiting the ability of the device to indicate relative delay values. By continuously sweeping the carrier over

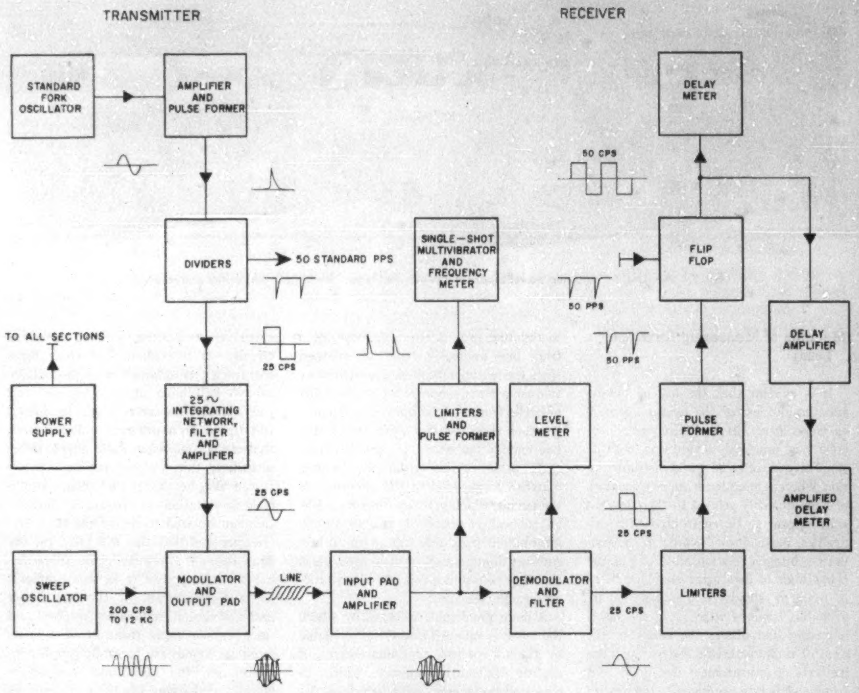

TRANSMITTER
RECEIVER
STANDARD FORK OSCILLATOR
AMPLIFIER AND PULSE FORMER
DIVIDERS
50 STANDARD PPS
TO ALL SECTIONS
POWER SUPPLY
25 ~ INTEGRATING NETWORK, FILTER AND AMPLIFIER
25 CPS
25 CPS
SWEEP OSCILLATOR
200 CPS TO 12 KC
MODULATOR AND OUTPUT PAD
LINE
INPUT PAD AND AMPLIFIER
LIMITERS AND PULSE FORMER
SINGLE—SHOT MULTIVIBRATOR AND FREQUENCY METER
LEVEL METER
DEMODULATOR AND FILTER
25 CPS
50 PPS
50 PPS
50 CPS
25 CPS
DELAY METER
FLIP FLOP
PULSE FORMER
LIMITERS
DELAY AMPLIFIER
AMPLIFIED DELAY METER

Fig. 2. Block diagram

the useful frequency range and measuring the frequency of the carrier at the receiving end, a plot of the delay information versus frequency is immediately and continuously available, which may be displayed on indicating meters or a cathode-ray oscilloscope.

Therefore, the 3-frequency method will permit measurement of envelope delay characteristics and, in addition, amplitude characteristics versus frequency. It will permit immediate and continuous display of the information with respect to frequency, the frequency axis being obtained by measurement of the carrier frequency as it is swept through the desired frequency band. Absolute measurements of delay can be made when both ends of the circuit are available at the same location. Relative values can be measured when the two ends are in isolated localities. Carrier-derived circuits in which the received frequency may differ slightly from the transmitted frequency, as well as physical circuits, may be measured.

Operating Principles

The envelope delay-measuring instrument embodies a practical application of the 3-frequency method of delay measurements as derived by an amplitude modulation process. It is a portable device suitable for the measurement of envelope delay and amplitude characteristics of circuits or equipment used for the transmission of information in the audio-frequency range. It consists of two units as shown in Fig. 1(A) and 1(B), each unit having over-all dimensions of $9^1/_2$ by $9^1/_2$ by $19^1/_4$ inches. Unit 1 consists of a regulated power supply and a sweep oscillator, and unit 2 a transmitter and receiver. Since much of the circuitry is common to both transmission and reception, these two functions were combined in a single unit.

Measurements may be made over a frequency range from 200 to approximately 12,000 cycles per second (cps). For loop measurements, only one instrument is required, in which case both absolute and relative delay as well as amplitude measurements can be made. In this case, the output of the transmitter is simply passed through the circuit whose characteristics are to be obtained, and thence to the input of the receiver. When the two ends of a circuit are in isolated localities, straightaway measurements are employed. In this case, an instrument is required at each end of the circuit, permitting relative delay and permitting amplitude measurements to be made in either direction.

Three methods are available for displaying the delay and amplitude characteristics with respect to frequency: 1. by means of the four meters on the panel of unit 2, 2. an external oscilloscope for which terminals are provided, and 3. external strip chart recording meters which may be inserted in any of the four meter circuits by means of the jacks provided. The delay and amplitude readings along with their associated frequencies appear on separate meters having delay scale ranges of 0 to 20, 0 to 5, and 0 to

Vol.
Pt. 1
1955

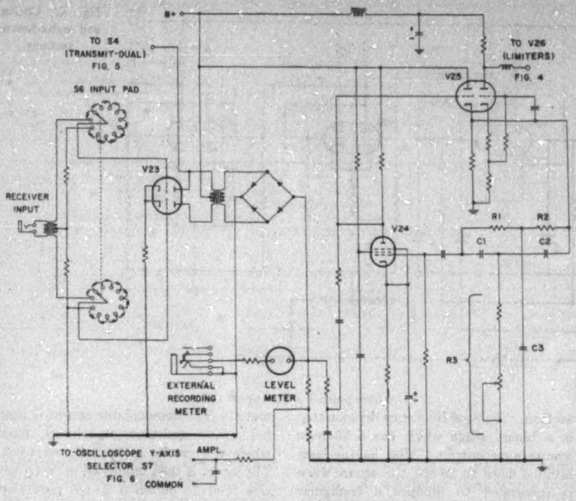


Fig. 3. Receiver input section

1 milliseconds, and frequency scales of 0 to 3, 0 to 6, and 0 to 12 kilocycles. For indicated delay greater than 20 milliseconds, the instruments continue to function but the proper multiple of 20 milliseconds must be added to the readings as determined by other known conditions of the circuit. The meter display is satisfactory for many purposes such as the indication of the flatness of delay and amplitude characteristics versus frequency in the application of corrective measures. Where permanent records of the characteristics are required, the meter display entails the almost simultaneous reading of two meters and recording the information while the frequency is being swept slowly, or it entails provision for the frequency to be varied in steps at the transmitting end. To make the information more readily obtainable, a cathode-ray oscilloscope display method was adopted. The horizontal axis is connected to the frequency-measuring circuit, and the vertical axis is connected to the delay or amplitude circuits. This arrangement makes immediately available a plot of relative delay or amplitude with respect to the sweep frequency being used at the transmitting end. This curve may be either photographed or traced and kept for permanent records.

The instrument operates basically by comparing the time relationship of two sets of pulses. One set is derived from an accurate tuning-fork standard oscillator; the other from the 25-cps envelope of a modulated wave that has been transmitted through the circuit whose characteristics are to be obtained. The 25-cycle modulating frequency is derived by frequency division from the standard oscillator. The carrier is generated by a separate sweep oscillator, the frequency of which is adjusted to cover the frequency band of the circuit under observation. Thus the time position of the second set of pulses with respect the standard set of pulses varies in accordance with the envelope delay of the circuit.

A block diagram of the instrument is shown in Fig. 2. In loop measurements where one instrument is used as both transmitter and receiver, a single standard oscillator is the source of both standard and comparison pulses. The time relationship between these pulses then depends only upon the envelope of the modulated wave transmitted over the circuit being tested, which is a function of the carrier frequency, and the additional delay which results from the instrument circuitry. This latter delay remains essentially constant over the frequency range of the instrument and may be subtracted from the indicated value to yield absolute delay. In straightaway measurements where a separate instrument is used at the transmitting and receiving ends, two different standard oscillators are supplying the standard and comparison pulses respectively. Therefore, the time relationship between them depends not only upon the criteria stated for loop measurements, but also upon the phase relationship between the two standard oscillator output waves and upon the slight frequency difference which may exist between the two standard oscillators. Consequently, it is practical to make measurements on a relative basis only, unless additional means are provided for phasing the standards at the two ends of the circuit. Since absolute values are very rarely required in straightaway measurements, it is not necessary to provide this means which is not easily accomplished. It is necessary only that the drift between the two standards be slow enough to permit time for the measurement to be made within the limits of allowable delay error. This drift can be made negligible for long periods by manual adjustment of the standard oscillator at the receiving end. Thus, slight differences between the standard frequencies at the two ends of the circuit do not incapacitate the instrument.

The circuitry of the instrument may be divided essentially into eight sections; the regulated power supply and sweep oscillator contained in unit 1, Fig. 1(A), and the standard oscillator, transmitter, receiver input, frequency-indicating, clipping and pulse-forming and delay-indicating sections contained in unit 2, Fig. 1(B). The power supply is conventional in design for supplying the plate and heater potentials.

Sweep Oscillator

The sweep frequency oscillator is of the beat frequency type with two high-frequency oscillators operating at about 135 kilocycles. The output of each oscillator is fed to separate cathode followers, the outputs of which are impressed upon the grids of a twin-triode grid bias modulator. A low-pass filter removes the carrier- and high-frequency modulation products but permits the passage of the difference frequency. Since the two triodes operate in push-pull, even harmonics of the difference frequency are suppressed in the output when the modulator is properly balanced. The modulator output is applied to a twin-triode output stage which includes additional filtering. The sweep frequency then is a sine wave having harmonic content below −45 decibels (db).

Continuous sweeping of the frequency is performed by a variable air capacitor in the tuned circuit of one of the oscillators, while the frequency of the other remains fixed. The variable capacitor is mechan-

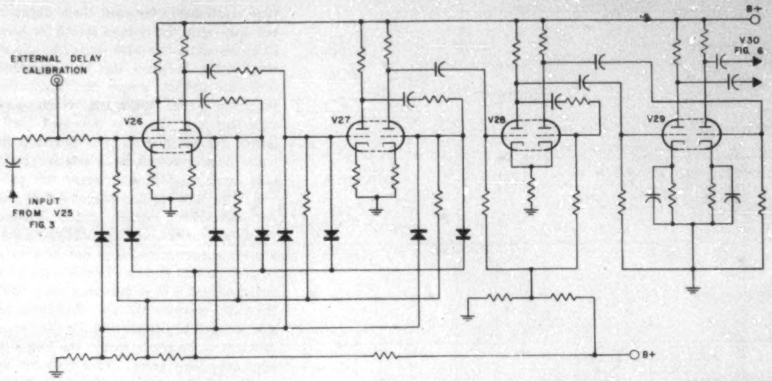


Fig. 4. Limiter and pulse-former sections

ically driven by a synchronous motor so as to provide a constant repetition rate for display purposes. The sweep repetition rate is varied in four steps by means of a gear system controlled by a rotary lever on the front panel of unit 1. This capacitor is in parallel with one of much larger value than itself and is so constructed that the spacing between the stator and rotor plates may be varied. This arrangement provides a means of varying the band of frequencies covered by the sweep range while maintaining a linear-frequency sweep throughout the range. The other oscillator contains a variable capacitor in its tuned circuit which permits the setting of the sweep-band of frequencies at a particular position in the spectrum and also permits setting the sweep-band limits without a trial-and-error process.

Standard Oscillator

The standard oscillator consists of a precision tuning-fork unit and a twin-triode amplifier stage. The output of this stage is coupled through a small capacitor to a heavily biased class-*B* amplifier which acts to suppress the negative swings of the oscillator waves and produce steep positive pulses.

Transmitter

The positive pulses from the standard oscillator section are used to control a series of cathode-driven frequency dividers which produce first negative pulses of very short duration having a repetition rate of 50 cps. The 50-cps pulses are used as the standard reference from which the delay indication is derived as discussed further in the delay-indicating section. The final frequency division step is a binary stage which has a 25-cycle square-wave output. After having been divided down to 25 cps, the square wave is integrated to produce a triangular wave shape which may then be readily converted to a pure sinusoidal wave shape by a tuned network and a push-pull amplifier.

The 25-cps output from this amplifier and the sweep frequency from the sweep oscillator of unit 1 are transformer coupled to a balanced modulator which suppresses the 25-cps modulating frequency in the output. This modulator is of a double pentode grid-bias type using remote cutoff tubes with balanced transformer output accurately adjusted to 600 ohms. The output signal is approximately 50 per cent amplitude-modulated having distortion components not in excess of −45 db with respect to the carrier level and practically constant level throughout the frequency band. The modulated signal is fed to the circuit to be tested through an adjustable output pad.

Receiver Input

The receiver input section, Fig. 3, is coupled to the line by a balanced transformer having an input impedance of 600 ohms. From the transformer, the signal is amplified in a twin-triode push-pull stage from which the signal is used to drive the frequency-indicating section as described in a following section. Also the output from the amplifier is used to drive a diode bridge demodulator, the output of which is first used to provide a source for the level meter, a voltage for oscilloscope display of input level versus frequency and also a source for use with a recording meter. The demodulator output is next fed to an resistance-capacitance filter where the carrier is partially removed. The signal is then impressed upon an active filter to obtain a 25-cps pure sine wave. The filter consists of a high-gain pentode amplifier, a cathode follower, a parallel-T blocking network in a feedback path and an output amplifier. This system not only removes the sweep frequencies, but also harmonics of 25 cycles, d-c variations, and any 60 cycles present. In addition, the system has a very important property in that unwanted delay variations due to rapid changes in signal level are opposite to the variations produced by the previous resistance-capacitance filter. The parallel-T network can be adjusted by means of R_3 so that practically perfect cancellation results.

Limiter and Pulse Former

The recovered 25-cycle waves are fed to a series of four diode limiter stages as shown in Fig. 4. The clipped waves are then fed to a stage of grid clipping after which the waves are differentiated and impressed on a class-*B* stage so that only negative pulses appear in the output from the plates and the positive pulses are partially suppressed. A double diode further removes the positive pulses and because of the phase relationship produces 50 negative pulses of very short duration from the original 25-cycle waves.

Frequency Indicating

The frequency-indicating section may be used to measure the frequency of either the transmitted or received signal, depending upon the manner in which the

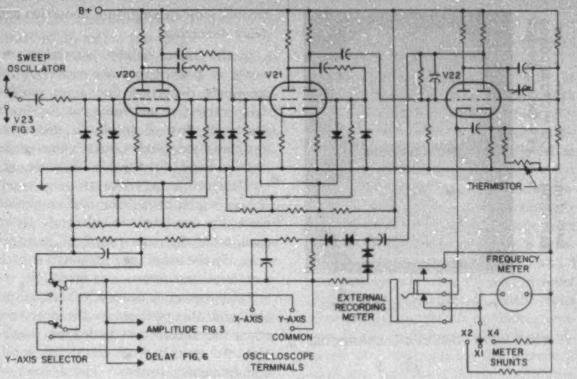

B+
SWEEP OSCILLATOR
V23 FIG 3
V20
V21
V22
THERMISTOR
FREQUENCY METER
EXTERNAL RECORDING METER
X-AXIS
Y-AXIS
COMMON
OSCILLOSCOPE TERMINALS
AMPLITUDE FIG 3
DELAY FIG 6
Y-AXIS SELECTOR
X2
X4
X1
METER SHUNTS

Fig. 5. Frequency-indicating section

instrument is being used. The frequency of the amplified signal of the input section or the frequency of the sweep oscillator at the transmitting end is impressed upon four limiter stages, Fig. 5, which produces a short pulse for each carrier cycle. These pulses are then used to trip a 1-shot multivibrator once for each cycle of the carrier, the output of which is averaged in the meter circuit to give an accurate indication of transmitted or received frequency. In addition, a separate voltage proportional to frequency is derived from the plate circuit of the multivibrator for application to the *X*-axis of an oscilloscope to display delay or amplitude characteristics versus frequency.

Delay Indicating

The delay-indicating section, Fig. 6, employs a flip-flop circuit in which a negative pulse in one grid trips one plate on and the other off, and a negative pulse in the other grid trips the flip-flop tube in the opposite direction so that only one-half of the tube may be on at any time. The tripping is done with two sets of negative pulses at a 50-cycle rate, one set being applied to each grid. The standard set is derived from the 50-cycle divider stage as described previously. The other set is derived from the recovered 25-cycle waves in the clipping and pulse-forming section. Since both sets of pulses are of the same repetition rate, the average current in each half of the flip-flop tube depends upon the time relationship of the two sets of pulses. A measure of the average current difference between the two halves of the tubes is then an indication of envelope delay of the circuit being tested. An additional delay results from the passage of signal currents through the circuits of the instrument. This delay, however, remains practically constant with frequency and may be subtracted from the indicated values when absolute values are required in the case of loop measurements.

A d-c meter is bridged between the cathodes to measure average current differences. Since two pulses per cycle are derived from the 25-cycle waves, the largest scale is 0 to 20 milliseconds. In addition to this scale, there are seven scales, each of which has a range of 5 milliseconds. These latter scales are located in successive positions throughout the 20-millisecond range with 2.5-millisecond overlap between adjacent scales, and are obtained by a d-c bias which simply shifts the scale but does not change its delay range. When either extreme of the 0-to-20-millisecond range is reached, the meter shifts abruptly to the opposite extreme so that operation at these points is not feasible. A means has been provided in the instrument for shifting to a more favorable position in the 0-to-20 range.

The voltage on one plate of the flip-flop is amplified by a long time constant resistance-capacitance stage to provide a 0-to-1-millisecond scale. Inverse feedback in the cathode further lengthens the time constant. A meter between the

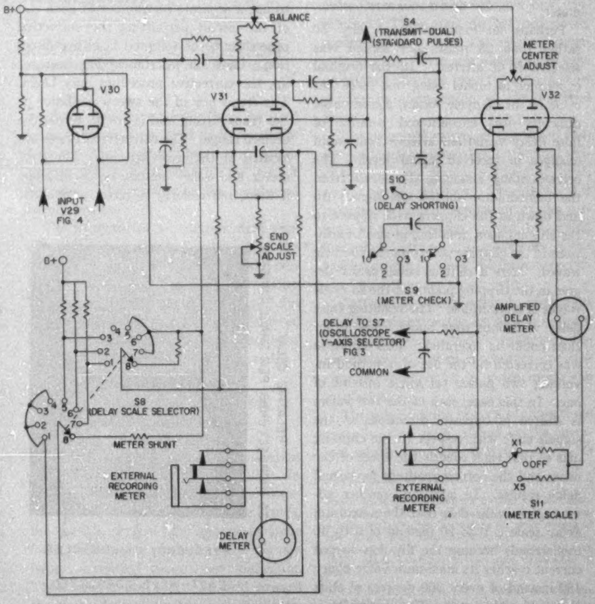

B+
BALANCE
S4 (TRANSMIT-DUAL) (STANDARD PULSES)
METER CENTER ADJUST
V30
V31
V32
INPUT V29 FIG 4
S10 (DELAY SHORTING)
END SCALE ADJUST
B+
S9 (METER CHECK)
AMPLIFIED DELAY METER
DELAY TO S7 (OSCILLOSCOPE Y-AXIS SELECTOR) FIG 3
COMMON
S8 (DELAY SCALE SELECTOR)
METER SHUNT
EXTERNAL RECORDING METER
DELAY METER
EXTERNAL RECORDING METER
X1
OFF
X5
S11 (METER SCALE)

Fig. 6. Delay-indicating section

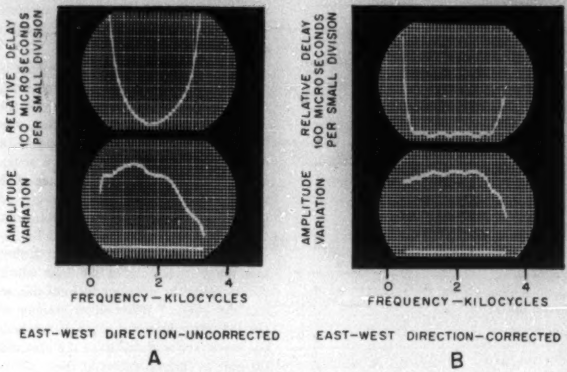


Fig. 7. Relative delay and amplitude oscillograms, 5-section type-K

A—Uncorrected
B—Corrected

cathodes of this stage provides an indication of relative delay. Also, the voltage on one of the cathodes provides the oscilloscope display of relative delay. Since the oscilloscope gain is adjustable over a wide range, it may be used to indicate a greater range of delay characteristics than the meter.

Perhaps an explanation of why 50 rather than 25 pulses per second was selected is of interest. In the original experimental model using one pulse per cycle of the 25-cycle waves, considerable difficulty was encountered because of false delay variations arising from rapid changes in received signal level. The corresponding changing d-c levels from the demodulator shifted the signals up and down in the clippers with respect to the clipping axis, resulting in small variations in the clipping points on the 25-cycle waves. Here a shift as small as 0.1 degree in the clipping points on the 25-cycle waves is undesirable. The resulting false delay variations were many times larger than could be tolerated. This situation was corrected by the use of a method involving two pulses per cycle instead of one. In this case, each of the two pulses is shifted in opposite directions as the signals shift with respect to the clipping axis so that each cancels the effect of the other and the correct readings for actual delay results. In addition, another advantage is realized in that the maximum delay scale is 0 to 20 instead of 0 to 40 milliseconds because the flip-flop output current reaches its maximum value every 180 instead of every 360 degrees of shift in the 25-cycle waves. The sensitivity is thus doubled.

Performance

The instrument was developed primarily for use in conjunction with corrective networks for equalization of delay and amplitude characteristics of facsimile circuits. For this purpose it is necessary to use a separate instrument at each end of the circuit permitting the corrective procedure to be followed in either direction. Once the instrument has been set up, the corrective procedure may begin. The frequency of the sweep oscillator at the transmitting end is swept across the desired range. The corrective devices are located at the receiving end. The level meter and delay meters, or oscilloscope if used, particularly to make permanent records, indicate amplitude distortion and delay distortion.

In the use of passive corrective networks, it is usual to make correction for amplitude distortion first since passive delay corrective networks can be made practically free of amplitude distortion but purely amplitude corrective networks possess delay distortion. However, active corrective networks are now available in which both delay and amplitude correction can be made independently of each other by adjustment of resistors only. In the use of the instrument as each new corrective insertion is made, the remaining correction necessary for equalization and the position in the spectrum where the correction is needed may be determined quickly from the delay, level, and frequency meters. The process of observation and correction is continued until the delay and amplitude variations are within prescribed limits.

The instrument is at present designed to cover a frequency range of approximately 200 to 12,000 cycles. The sweep oscillator can be set to sweep this entire range or any portion of this band not less than about 900 cycles in width. Over this range the delay and amplitude characteristics due to the instrument circuitry are essentially constant. The transmitter output level to the line is adjustable in five steps from +10 decibels above 1 milliwatt to −10 decibels below 1 milliwatt. The receiver will operate stably for constant input levels from the line over a range from approximately −35 decibels below 1 milliwatt to +20 decibels above 1 milliwatt with appropriate pad settings. With a varying input level from the line and a fixed pad setting, the re-

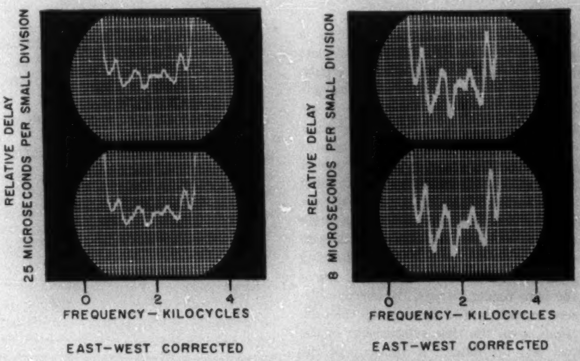


Fig. 8. Relative delay oscillograms, 5-section type-K, corrected

ceiver will operate over a variation of at least 30 db within this range.

The accuracy of the instrument depends largely upon the means used to display the information, the type of measurements desired, and the characteristics of the circuits or equipment required to be measured. Under favorable conditions, particularly those in which the delay characteristics are relatively flat throughout the frequency band under consideration, an accuracy of approximately ±10 microseconds may be attained. For conditions in which large delay variations are encountered, the accuracy attained is considerably less, extending to approximately ±50 microseconds. For external displays, such as an oscilloscope display, the accuracy depends upon the precision with which the frequency, delay, and amplitude axis settings are made. In any case, the delay information can be checked by the insertion of known delay values for which one means is provided in the instrument. Typical oscilloscope displays are shown in Figs. 7 and 8.

These oscillograms were made with the use of a long 5-section type-*K* circuit both uncorrected and corrected, in which only a partial correction was made for amplitude distortion. The delay distortion correction is effective in producing a flat characteristic throughout the useful frequency band to within approximately ±100 microseconds. Two records were made on each oscillogram of Fig. 8 at expanded scales of approximately 25 and 8 microseconds per small division in order to determine the effect of circuit noise occurring in long circuits on the performance of the instrument. It will be observed that the resulting disturbance is small and that the departure in relative delay between these two traces within the useful frequency band does not exceed approximately ±10 microseconds.

Aside from its use in the correction of delay and amplitude distortions in which the instrument greatly exceeds the accuracy requirements, it is finding use in other applications where greater accuracy is required. In these applications, the instrument may be used to measure and make permanent records of the characteristics by the sweep method or by point-by-point methods, particularly where greater accuracy is required for absolute measurements on a loop basis.

References

1. The Building Up of Sinusoidal Currents in Long Periodically Loaded Lines, John R. Carson. *Bell System Technical Journal*, New York, N. Y., Oct. 1924.

2. Phase Distortion and Phase Distortion Correction, Sallie P. Meade. *Ibid.*, May 1928.

3. Measurement of Phase Distortion, H. Nyquist, S. Brand. *Ibid.*, May 1930.

4. Phase and Amplitude Distortion in Linear Networks, M. J. DiToro. *Proceedings*, Institute of Radio Engineers, New York, N. Y., Jan. 1948.

The Effect of Reactor Irradiation on Electrical Insulation

J. C. PIGG
NONMEMBER AIEE

C. D. BOPP
NONMEMBER AIEE

O. SISMAN
NONMEMBER AIEE

C. C. ROBINSON
NONMEMBER AIEE

Synopsis: A brief review of the usual causes of insulation deterioration is given. The effect of radiation on these processes is discussed. It is observed that additional factors are present when an insulation is in a radiation field which change both the quality and nature of the insulation.

INSULATION in its broadest sense implies a method of isolation of a system. In electric systems, insulation may also serve as part of a component as in the case of a dielectric; however, for the present purpose, insulation will be considered to be a means of isolating an electric system.

Adequate insulation may be defined as that degree of isolation which holds spurious circuits below the maximum value permitted by engineering tolerances. The requirements for adequate insulation vary radically as the nature of the circuit and its environment are changed. For example, a short-circuited stub may be an open circuit at one frequency, an inductance or capacitance at another frequency, or a short circuit at some other frequency and direct current. Conversely, a polystyrene rod which is an insulator for direct current may be used as an antenna element at microwave frequency. The voltage, frequency, application, and environment of the circuit determine the nature of the insulation requirements.

The presence of a radiation field produces a new and drastic element in the environmental factor which alters the performance of insulation both in quality and nature. Of these two, perhaps change in quality should be considered first since it is a natural extrapolation of normal experience with such materials.

Changes in Quality of Insulation in Normal Use

A certain amount of ionic conduction occurs in all insulating materials.[1] In many inorganic materials, while the skeleton of the structure has homopolar bonds, the metals present are mainly in ionic form. Examples of such materials are mica, asbestos, glass, and porcelain. These ions migrate in a d-c field toward the electrodes and produce a slow disintegration of the material. In ionic materials, crystal imperfections give rise to vacant lattice sites which migrate through the crystal producing a current. In organic solids, dissociation of neutral molecules forms ions which can give ionic conduction. These ions do not necessarily recombine to reform the original molecule, but may form new molecules or take up water or oxygen from the air. A slow change in composition of the material results.

The principal causes of insulation failure are surface breakdown, external discharge, internal discharge, moisture, heat, and aging. Surface breakdown is brought on by contamination of the surface or deterioration caused by arcing from high voltages. The presence of the insulation may reduce the flashover voltage below that for an air path.

External discharge refers to discharges which impinge on the insulation and are sustained by currents flowing through the insulation. Internal discharges are discharges which occur in voids inside the insulation.

The deterioration in each of these three cases is sometimes caused by the chemical attack from the activity of oxides

---

Paper 55-694, recommended by the AIEE Dielectrics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted June 10, 1955; made available for printing September 7, 1955.

J. C. Pigg, C. D. Bopp, and O. Sisman are with the Oak Ridge National Laboratory, Oak Ridge, Tenn., and C. C. Robinson is with the Wright Air Development Center, Dayton, Ohio.

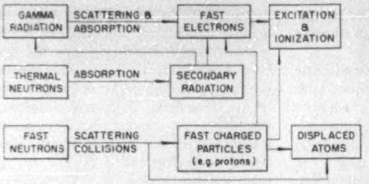

GAMMA RADIATION
SCATTERING & ABSORPTION
FAST ELECTRONS
EXCITATION & IONIZATION
THERMAL NEUTRONS
ABSORPTION
SECONDARY RADIATION
FAST NEUTRONS
SCATTERING COLLISIONS
FAST CHARGED PARTICLES (e.g. protons)
DISPLACED ATOMS

Fig. 1 (above). Prominent reactions of reactor radiation with matter

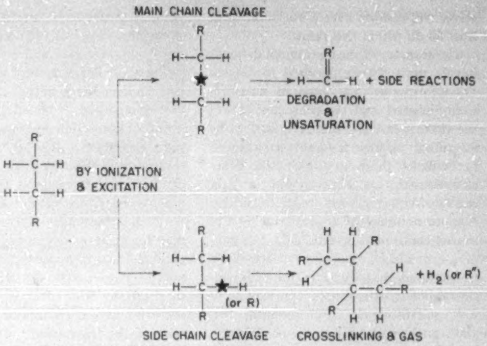

MAIN CHAIN CLEAVAGE
+ SIDE REACTIONS
DEGRADATION & UNSATURATION
BY IONIZATION & EXCITATION
(or R)
$+H_2$ (or R″)
SIDE CHAIN CLEAVAGE
CROSSLINKING & GAS

Fig. 2 (right). Reactions of ionizing radiation on a hydrocarbon polymer

of nitrogen and ozone, although deterioration of many materials is as rapid in the absence of oxygen as in its presence. These effects are more pronounced in the organic insulations. Polymerized hydrocarbons are more resistant than most other organic insulations. The replacement of oxygen by nitrogen, hydrogen by halogens, and carbon by silicon gives an additional increase in resistance to this type of breakdown.

The presence of moisture can increase the leakage current caused by conduction on moist surfaces; it can penetrate the structure of the insulation where the large electric dipole moment of the water molecule can change the dielectric properties of the insulation and it can, to some extent, dissolve parts of the insulation, thus breaking it up into ions. Substances which have high electrical resistivity in general have a high resistance to heat flow. Heat generated within the insulation caused by d-c leakage currents or by a-c power-factor losses will raise the temperature of the material until the loss of heat by conduction equals the rate of introduction. The electrical conductivity of an insulator increases with increase in temperature which further increases the rate of generation of heat. The rate of deterioration caused by the process just mentioned also increases with temperature. Increase in temperature can cause the material to be mechanically unstable.

Aging is a general term applied to any slow deterioration the mechanism of which has not been satisfactorily explained, or to an aggregate of effects when a general term is convenient. Organic compounds can be expected to undergo chemical changes which will alter their electrical behavior. Surface contamination, absorption of water and other contaminants from the atmosphere, and mechanical attacks on the material integrity accumulate in time to change the effectiveness of system isolation.

Changes in the Quality of Insulation Under Reactor Irradiation

RADIATION REACTIONS

The change in the mechanical and electrical properties of insulating type materials have been studied extensively.[2,3] Much of this information is the result of before- and after-type measurements and usually slabs of material are used. The data reported here are essentially dynamic in that measurements are made during exposure. In addition, the specimen geometry is not idealized.

The presence of radiation may accelerate the deteriorating mechanisms mentioned in the foregoing and introduce new considerations. Since ionization is produced by high-energy radiation in the material and surroundings, flashover and breakdown voltages are reduced. Leakage currents are increased.[4] The material is internally heated by absorbing energy from the radiation field. Chemical bonds are broken both from Compton scattering by gamma rays and mechanically by neutron collision and the recoil particle ionization. Ozone and other active ions are formed and may attack the material. Compounds of nitrogen are formed both from atmospheric nitrogen and nitrogen in the material. The material may become more hygroscopic which increases all the difficulties from absorbed moisture.[5] New impurities are produced by transmutation in the bulk of the material. The net effect so

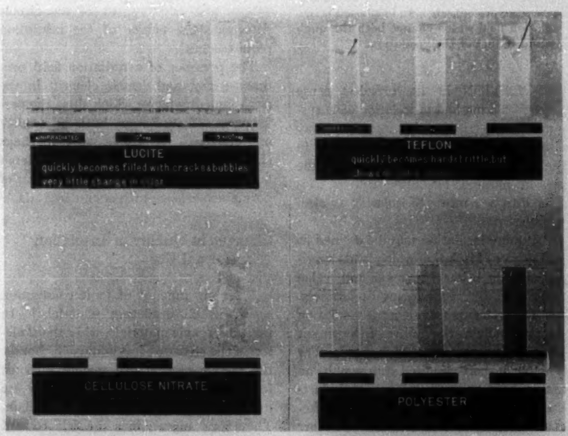

UNIRRADIATED
LUCITE
quickly becomes filled with cracks & bubbles
very little change in color
TEFLON
quickly becomes hard & brittle but
CELLULOSE NITRATE
POLYESTER

Fig. 3. Changes produced by ionizing radiation in lucite, Teflon, cellulose nitrate, and polyester

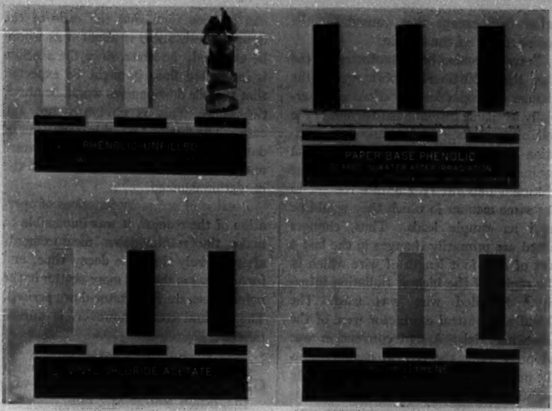


Fig. 4. Changes produced by ionizing radiation in unfilled phenolic, paper-based phenolic, vinyl chloride acetate, and polystyrene

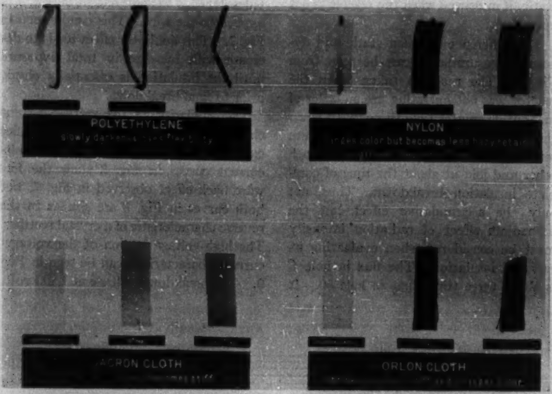

POLYETHYLENE
slowly darkens, loses flexibility
NYLON
changes color but becomes less hazy, retains
DACRON CLOTH
ORLON CLOTH

Fig. 5. Changes produced by ionizing radiation in polyethylene, nylon, dacron, and orlon

Table I. Changes in Bulk Properties Caused by Radiation

Material	Volume Resistivity, Ohm-Centimeters, for Radiation Exposures of 0	5×10^8	5×10^9	10^{10}	Physical Change
Urea formaldehyde	2×10^{13}		10^{11}		cracking
Cellulose acetate	5×10^{13}	10^{11}			cracking
Polyethylene	$>10^{14}$			$>10^{14}$	
Nylon	10^{13}		4×10^{13}	$>>4\times10^{13}$	hardening, water absorption
Phenolic, asbestos filler	2×10^9		10^{10}		hardening
Phenolic, paper filler	3×10^{11}		10^{12}	10^{11}	hardening, degradation, swelling
Polyester, CR-39	$>10^{14}$		10^{13}		cracking, decomposition products
Polyester, selectron	10^{11}		10^{14}	$>>10^{14}$	hardening, cracking
Polystyrene	$>10^{14}$			$>10^{14}$	
Vinyl chloride acetate	$>10^{14}$	$>10^8$			decomposition product HCl, water absorption

far discussed, with regard to the structure of the insulator, is that of accelerated aging.

Inorganic materials are ionized by gamma radiation and heated by energy absorbed from the radiation field. Both effects lead to increased conductivity. In addition, disruption of the crystal structure and the introduction of impurities by fast and slow neutrons respectively result in permanently lowered resistance. Ionic materials experience the same effects with the added consideration that a permanent change is introduced by the increased density of lattice defects.[5] Organic materials, while subject to the same mechanisms operating in the inorganic and ionic materials, are changed primarily because of ionization caused by gamma rays and high-energy neutrons in the reactor.

The effect of high-energy radiation on a polymer molecule is the breaking bonds. Recombination may occur at the site of the broken bond—with no net change—or the ions may diffuse away before they can recombine. When the site of the broken bond is in a side chain of a long chain polymer, two such molecules may join to form one large molecule to a network-type structure. This mechanism is called crosslinking. When the broken bond occurs in the main chain and results in splitting the molecule, the mechanism is called chain cleavage. Both crosslinking and chain cleavage are often accompanied by the evolution of gas.

Slow neutrons may be captured and thus produce secondary radiation and small amounts of impurities. Fast neutrons are slowed down principally by collision with hydrogen nuclei causing recoil protons which, being charged particles, cause excitation and ionization along the path of the recoil ion; see Figs. 1 and 2. Usually the effects of thermal neutron irradiation on polymers is small compared to that of the gamma ray and fast neutron irradiation. However, the effect of thermal neutron irradiation may become appreciable in cases where atoms having high cross sections are present in the polymer.

Changes were suggested in the foregoing which improved the resistance to decay under normal conditions. These changes in some cases reduce the resistance to deterioration by radiation. Halogens, for example, produce acids which attack the material as well as increase the bulk and surface conductivity. The presence of chlorine with its large capture cross section renders any polymer containing it particularly vulnerable to damage by slow neutrons.

Table II. Breakdown Dosages of Several Insulators

Material	Dosage, Neutrons per Centimeter²	Comment
Polyethylene	10^{19}	still good on removal
Silastic 80	10^{19}	still good on removal
Sil-X	9×10^{18}	breakdown
Teflon	5×10^{18}	breakdown
Silicon rubber	4×10^{18}	breakdown
Neoprene	3×10^{18}	breakdown
Formvar	2×10^{18}	mechanical failure
Polyvinyl chloride	1.9×10^{18}	breakdown
Glass cord covered with glass braid	1.9×10^{18}	breakdown
Rubber	1.3×10^{18}	breakdown
Kel *F*	10^{18}	breakdown
Surprenant *A-10*	10^{17}	breakdown
Surprenant *B-2*	5×10^{16}	breakdown

PHYSICAL CHANGES OBSERVED CAUSED BY RADIATION

The changes observed in a polymer depend upon the relative rate of crosslinking and chain cleavage. Some results of chain cleavage are:

1. Softening.
2. Lower tensile strength.
3. Increase in elongation.
4. Lower shear strength.
5. Lower melting point.
6. Greater solubility.

Some results of crosslinking are as follows:

1. Hardening.
2. Increased strength.
3. Increase in Young's modulus.
4. Increase in melting point.
5. Decrease in solubility.
6. Increase in density.

Some typical examples are shown in Figs. 3 through 5.

RESISTANCE CHANGE UNDER RADIATION

Some typical changes in the volume resistivity of several irradiated materials are shown in Table I. The resistivity was measured in the absence of any radiation field except that caused by the radioactivity of the sample.

Tests have also been conducted at the Oak Ridge National Laboratory on the change of properties of various commercially available insulations to evaluate the reliability of their performance in a radiation field.[7] The tests were conducted by installing the wires in the Oak Ridge National Laboratory graphite reactor in the same manner in which they would be used as sample leads. Thus, changes noted are primarily changes in the last 4 feet of a 22-foot length of wire which is the region of the highest radiation intensity.[8] Shielded wire was used. The shield and central conductor were of the same material. A sign convention was adopted using the central conductor as the positive terminal and the shield as the negative terminal. In multiple concentric configurations, the outermost conductor was taken as negative for any given measurement. The measurements are both dynamic, if measured when the reactor is operating, and they are static, if measured when the reactor is not operating.

The difference between static and dynamic measurements can be seen from Fig. 6. The readings taken when the reactor was not operating are an order of magnitude greater than those taken when the reactor is operating. As the exposure increases, the two curves approach each other and join at about the time of complete insulation breakdown. Thus, not only the accumulative effect but the immediate effect of radiation intensity must be considered when evaluating irradiated insulation. The flux in hole *C* is about three times that in hole 50. It can be seen then that the ratio of the dynamic resistances is about 3 as would be expected; however, since the abscissa is integrated flux, it might be expected that the pile down curves would be about the same. This is obviously not the case. The flux or rate of introduction of damage as well as the total exposure must be considered in evaluation of the insulation damage.

Because of necessary functions of operation of the reactor, it was impossible to make the reactor-down measurements after precisely the same decay time, and for this reason there is more scatter in the points taken during reactor-down periods. The general trend, however, is unmistakable.

Observed Changes in the Nature of Insulation Behavior Under Irradiation

RECTIFICATION

The resistance which the insulation provides depends to some extent on the bias applied to the wire. This can be seen in Fig. 7. This nonlinear effect tends to decrease with increase in total exposure until the insulation is essentially ohmic prior to final breakdown.

The similarity of the voltage-current characteristics to that of a crystal rectifier can be seen in Figs. 8 and 9. The volt-current curves in Fig. 8 show the forward-back effect observed in Fig. 7, but both curves in Fig. 8 are similar to the reverse characteristic of a crystal rectifier. The high-voltage section of the voltage-current characteristic can be seen in Fig. 9. The breakdown voltage at 1,400 volts

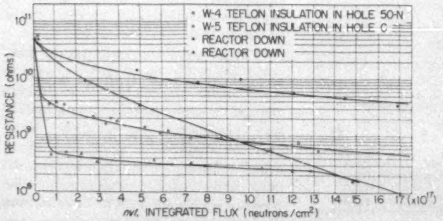


Fig. 6. Insulation resistance versus integrated neutron flux for different exposure levels

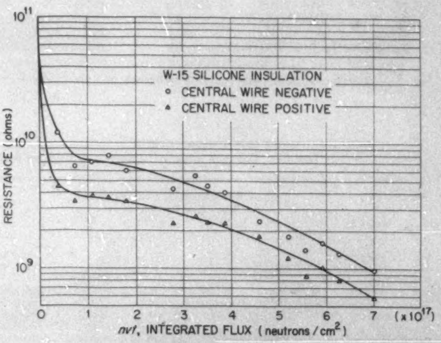


Fig. 7. Insulation resistance versus integrated neutron flux for different bias polarity

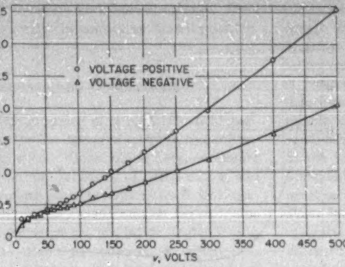


Fig. 8 (left). Applied voltage versus sample current for bombarded Teflon insulation

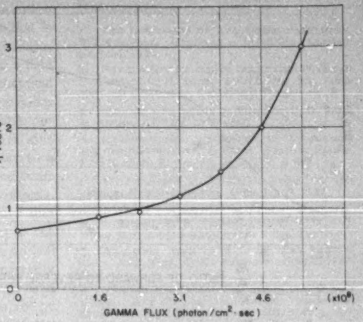


Fig. 10 (right). Photovoltage versus gamma flux for polyvinyl insulation

does not permanently impair the insulation but rather is indicative of a reversible effect. The curve has a slope of 7.5 megohms and the initial low-voltage behavior is recovered when the applied voltage is reduced.

PHOTOELECTRIC EFFECT

A voltage is observed on most insulated wires which are exposed to a gamma flux. A plot of observed voltage output versus gamma flux is shown for a polyvinyl chloride sample in Fig. 10. This sample was located in the reactor and thus, even at zero reactor power, it still was exposed to a considerable gamma field. This is seen in the figure as a finite voltage at zero flux. The photovoltage is related to both the accumulated bombardment and the ratio of forward-to-reverse resistance, as shown in Figs. 11 and 12 for a silicone sample.

NONOHMIC BEHAVIOR

In discussing the nonohmic behavior wire insulation, an analogy with a crystalline solid is used. The electrons in an atom must occupy certain discrete energy states. When atoms are brought together to form a solid, these energy states form bands of energy levels, some of which are filled with electrons, some of which are empty, and some of which may be partially filled.

If the highest energy band containing electrons is partially filled or if the highest filled energy band overlaps the next band which is empty, then the electrons in such a band may be accelerated by an electric field and the material is a conductor. If the highest energy band containing electrons is filled and the next band is separated from it, then electrons cannot be accelerated and the material will not conduct electricity. If the gap between the energy bands is small enough, electrons may be thermally excited from the filled band to the empty band and conduction can occur. Such materials are called semiconductors. When the gap between the energy bands is so large that electrons cannot be excited from the filled band to the empty band in detectable quantities, conduction does not occur. Such a material is an insulator. The excitation of electrons across the energy gap by photons or the introduction of energy levels into the gap can produce semiconductorlike behavior in an insulator.

When a semiconductor and a conductor are placed in contact, a potential barrier is produced between them. Such a barrier produces rectification and photo voltages. An equivalent circuit for such a situation is shown in Fig. 13. The barrier is shown as a diode which has a series resistance of R_s. The diode and its series resistance are in parallel with the insulation resistance of the rest of the wire and any meter which may be applied for measurement. There are two barriers, one at the central conductor and one at the shield. As can be seen in Fig. 13, these two barriers are in opposite directions. It might then be expected that the voltage-current characteristics would be similar to those of the reverse direction of a crystal rectifier, such as was seen to be the case in Fig. 8. Since the contact areas are different at the two barriers, the shield-insulator barrier admits a larger current in the reverse direction than the conductor insulator barrier. This is seen

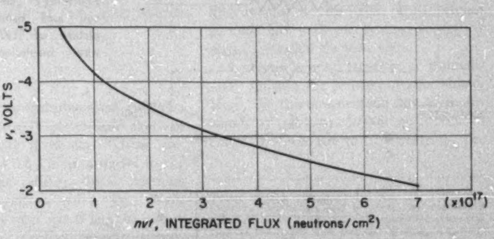


Fig. 11 (above). Photovoltage versus integrated neutron flux for silicone insulation

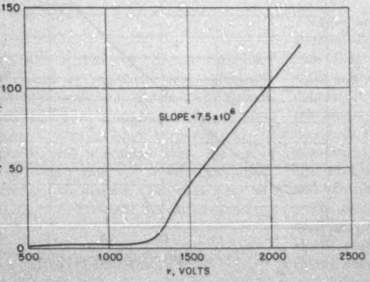


Fig. 9 (left). Applied voltage versus sample current for bombarded Teflon insulation

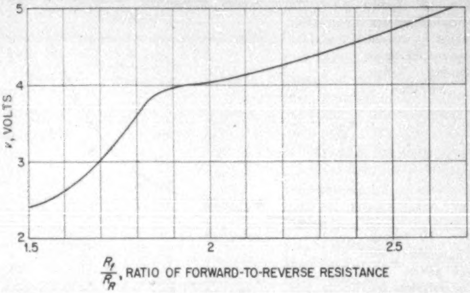


Fig. 12. Photovoltage versus resistance ratio for silicone insulation

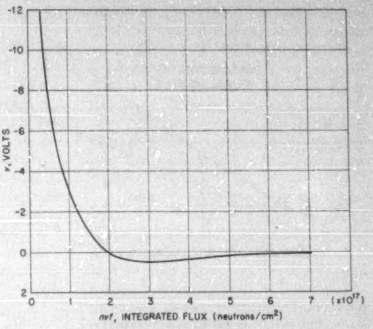


Fig. 14. Photovoltage versus integrated neutron flux for Teflon insulation

to be the case in Fig. 8. When the applied voltage becomes sufficiently larger than the barrier height, the resistance to the flow of current approaches R_s. This effect can be seen in Fig. 9. In this high-voltage region, both directions of bias yield the same current. At equilibrium, the total currents across the boundaries must equal zero.

Considering the phenomenon to be a diffusion process, the current density across the barrier can be written[9]

$$j=\frac{q\mu n}{\lambda}(V_b+V)\epsilon^{-\frac{qV_b}{kT}}\left(1-\epsilon^{-\frac{qV}{kT}}\right) \qquad (1)$$

where

j = current density
q = charge of the electron
μ = mobility of charge carrier
n = density of charge carriers in the material at a distance from the barrier
λ = barrier thickness
V_b = barrier height
V = voltage across the barrier
k = Boltzman constant
T = absolute temperature

The currents across the two barriers must be equal

$$I_1=I_2 \qquad (2)$$

$$a_1j_1=a_2j_2$$

where a is the area per unit length along the wire. Let r be the respective radius

$$2\pi r_1j_1=2\pi r_2j_2$$

$$j_1=\frac{r_2}{r_1}j_2 \qquad (3)$$

Combining equations 1 and 3 yields equation 4

$$(V_b+V_1)\left(1-\epsilon^{-\frac{qV_1}{kT}}\right)=\frac{r_2}{r_1}(V_b+V_2)\times\left(1-\epsilon^{-\frac{qV_2}{kT}}\right) \qquad (4)$$

Since $r_2>r_1$, the equation can hold only for $V_2<V_1$. The total voltage seen across the system will be $V=V_1-V_2$.

The bias direction of the high and low resistance, and the sign of the output voltage are determined by the sign of the charge carrier in the insulator. If conduction is made possible by electrons photoelectrically excited from the filled to the empty band, the material will be n-type. If conduction is caused by the presence of ionized atoms which can act as acceptors, the conduction will be due to holes and the material will be p-type. The change in voltage output for a Teflon sample as exposure continues is shown in Fig. 14. The negative voltage indicates that the sample is originally n-type, conduction being due to photoelectrons. As the exposure continues, an acceptor concentration builds up until the output voltage is positive, indicating the sample p-type. Fig. 15 shows the ratio of the

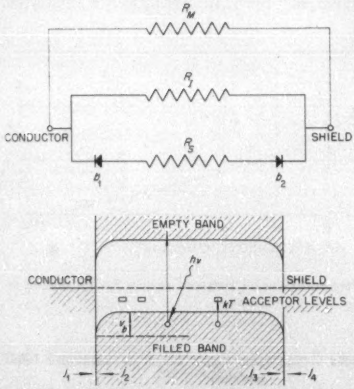


Fig. 13 (left). Equivalent circuit and band picture of lead wire insulation

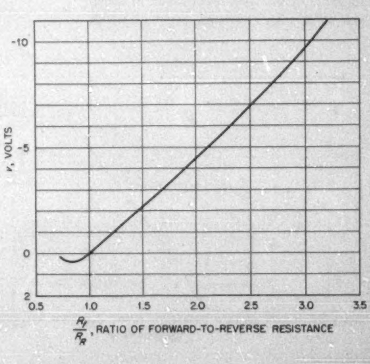


Fig. 15 (right). Photovoltage versus resistance ratio for Teflon insulation resistances for both bias directions. R_f is taken with the conductor biased positive.

Summary and Conclusions

The presence of a radiation field subjects the insulation of a circuit to environmental conditions which both accelerate the normal processes of deterioration and introduce new characteristics to the insulation behavior, as noted in Fig. 6. The resistance of an insulation under irradiation drops to some value which is a function of the flux level and maintains this value as irradiation proceeds. Eventually the insulation resistance begins to drop again. This final resistance drop can be designated the breakdown of the insulation. The breakdown irradiation dosage for several insulations studied is shown in Table II. The change in insulation resistance in the presence of a field is greater than that caused by chemical deterioration since the ionization provides mobile charges which are available for conduction. These charge carriers give to the insulation properties similar to a semiconductor and result in the nonohmic resistance as well as photovoltages. These photovoltages change in sign and magnitude depending on the sign of the charge carrier. Voltages as high as 200 volts have been observed in commercially available wires. The nonohmic and photovoltage effects can be minimized by proper geometrical arrangement of the insulation.

References

1. The Insulation of Electrical Equipment, **Willis Jackson.** Chapman and Hall Ltd., London, England, 1954.

2. Physical Properties of Irradiated Plastics, **O. Sisman, C. D. Bopp.** *ORNL-928*, Oak Ridge National Laboratory, Oak Ridge, Tenn.

3. Radiation Stability of Plastics, **C. D. Bopp, O. Sisman.** *ORNL-1373, Ibid.*

4. Recombination Processes in Insulators and Semiconductors, **A. Rose.** *Physical Review*, New York, N. Y., vol. 97, no. 2, 1955, p. 322.

5. *WADC Report WCRTY55-1*, Wright Air Development Center, Dayton, Ohio, Jan. 5, 1955.

6. Radiation Effects in Ionic Crystals, **R. L. Sproull, R. S. Caswell.** *ORNL-1359*, Oak Ridge National Laboratory, Oak Ridge, Tenn.

7. Radiation Effects in Lead Wire Insulation, **J. C. Pigg, C. C. Robinson.** *ORNL-1762, Ibid.*

8. Testing Electrical Insulation for Use in Gamma Ray Fields, **C. Mannal.** *Nucleonics*, New York, N. Y., vol. 12, no. 6, 1954, p. 49.

9. Metal Rectifiers, **H. K. Henisch.** Clarendon Press, Oxford, England, 1949.

---

l, 74
1
55

The Effects of High-Energy Gamma Radiation on Dielectric Solids

PHILIPP H. KLEIN
NONMEMBER AIEE

CLIFFORD MANNAL
ASSOCIATE MEMBER AIEE

Synopsis: The kinds and magnitudes of radiations present in the various locations near nuclear reactors are discussed as well as the mechanisms of interaction of radiation with matter, namely, group fragmentation, molecular scission, and polymerization. Techniques and locations for obtaining high dose rates are presented. Results of 60-cycle short-time voltage breakdown tests on polyethylene, cellulose acetate, and polyvinyl chloride films along with Formex*-varnished wire are given as a function of integrated dose up to 10^8 roentgens. Extensive electrical tests on a mica-glass-silicone-resin insulation simulating that used in the electromagnetic pump for the submarine intermediate reactor are described. Measurements of mechanical property changes of irradiated resins by means of a newly developed penetrometer are presented.

DURING the past decade, sizeable groups of engineers and scientists have been working to obtain useful power from the fission of uranium. Several designs for reactors have been reduced to practice, and a number more are very close to this state. It seems reasonable, therefore, to assume that each year in the next decade will see an increasing number of nuclear reactors in operation and that, at the end of this time, a small, but appreciable portion of the country's installed generating capacity will be powered by fissile materials.

It is the purpose of this paper to discuss: 1. the nature of nuclear radiations; 2. the locations and approximate intensities with which they exist in the vicinity of nuclear reactors; 3. the reaction processes of these radiations with matter; 4. experimental techniques; and 5. results of tests made at the Knolls Atomic Power Laboratory.

General

There are four fundamental particles abundantly present in nuclear reactors and their waste products. These are: alpha particles, i.e., helium nuclei in the doubly ionized state; beta particles (electrons); gamma rays, i.e., high-energy X rays with wave lengths corresponding to equivalent generating voltages of 1/2 to several millions of volts; and neutrons.

The range of charged particles in solid or liquid materials is very small. An alpha particle of 10-megelectron-volt (Mev) energy will penetrate a piece of steel only about 1 mil and an electron of the same energy will go only 0.66 inch. Except in gaseous media, alpha and beta radiations are thus largely absorbed by adjacent atoms. Appreciable fractions of the original neutron or gamma intensity may exist after travel through 1 foot or more of water, lead, or concrete. Table I lists the approximate magnitudes of neutron and gamma flux which are found in, near, or around nuclear reactors.

Within the reactor a thermal neutron flux of 10^{10} to 10^{12} neutrons per square centimeter per second or even higher can be expected. A fast neutron flux of perhaps 1/10th the magnitude of the thermal neutron flux will also be present. These fast neutrons may possess energies up to a number of million volts. The more energetic fast neutrons are very much more destructive than the thermal neutrons. Accompanying the high level of neutron flux in the reactor core is an intense gamma flux with an average energy of about 1,000,000 electron volts.

A second generic location is immediately outside the primary reactor shield. Here, the thermal neutron flux is diminished by perhaps 10,000 or 1,000,000 times, i.e., to 10^6 to 10^7 neutrons per

---

* Trade name of the General Electric Company for a polyvinyl formal insulation.

Paper 55-741, recommended by the AIEE Dielectrics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted June 7, 1955; made available for printing August 8, 1955.

Philipp H. Klein and Clifford Mannal are with the General Electric Company, Schenectady, N. Y.

The General Electric Company operates the Knolls Atomic Power Laboratory for the U. S. Atomic Energy Commission under Contract No. *W-31-109 Eng-52*.

Table I. Gamma and Neutron Fluxes in the Vicinity of Nuclear Reactors

Location	Neutron Flux, Neutrons per Centimeter²-Second	Gamma Flux, r/hr	Insulation Required for
Reactor cores	10^{12} to 10^{13} or higher	10^{8} to 10^{9}	instruments for in-pile measurements, e.g., thermocouples
Outside primary reactor shield	10^{5} to 10^{7}	10^{6} for Na coolant	control mechanisms
Inside secondary shield	10^{3} to 10^{7}	10^{4} for water coolant Negligible for He	pumps level indicators temperature indicators
Outside secondary shield	negligible	negligible	standard electric apparatus

square centimeter per second. It is customary to place some of the control and pumping equipment at this location. An essential difference is observed between the two most commonly used coolants at this point. Liquid sodium, which has many advantages from a heat-transfer standpoint, becomes intensely radioactive in the nuclear reactor. It emits gamma rays of 1.39 and 2.76 Mev. If sodium is used as a coolant in the reactor, the gamma flux in the vicinity of the sodium storage and standby tanks may be of the order of 10^6 roentgens per hour (r/hr). (By action of the International Commission on Radiological Units in July 1953, the roentgen is defined as "the quantity of X- or gamma-radiation such that the associated corpuscular emission per 0.001293 gram of air produces, in air, ions carrying 1 electrostatic unit of quantity of electricity of either sign." Essentially, this means 1 electrostatic unit per cubic centimeter of dry air at 0 degree centigrade (C) and 760 millimeters of mercury, or 83 ergs/gm. In water, one roentgen is equivalent to about 94 ergs/gm.) If the reactor is water-cooled, the radiation level will depend on the attenuation of the reactor shield. The figure of 10^4 r/hr might be a typical order-of-magnitude figure under this condition.

The third zone of application is outside the secondary or biological shield. Here, fluxes are sufficiently low so that personnel can tolerate them for extended periods without damage. No special precautions need therefore be taken with electric equipment in these areas.

The situation with respect to chemical processing of expended fuel or waste products is somewhat simpler. Here the neutron level is extremely low. Alphas and betas, as previously explained, are largely self-absorbed in the solvent materials. If large tanks of the materials are processed, substantial gamma-ray levels may influence the location of pumps, contactors, and controllers.

While a very few special cases of favorable reactions of nuclear radiations on organic materials have been reported, by far the most common effect is the deterioration of the physical, electrical, and mechanical properties of the materials.[1-5] The general process is very complex. In any given case, the effects caused by the various mechanisms may differ in degree, and may give rise to very different end products. Fundamentally, the first step in any of these reactions is the concentration of energy at a particular location in the molecule. After this one or more of the following processes may take place:

1. Group Fragmentation: By this is meant the "chipping off" of individual hydrogen, carbon dioxide, or similar small groups from a long and complicated molecule. A continued gas evolution is observed which, if it is continued in small sealed containers, can generate very high pressures.

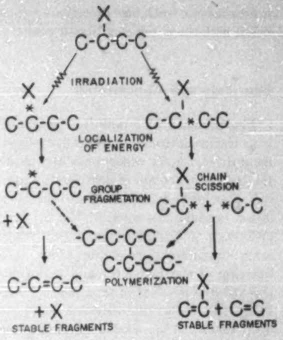


Fig. 1. Schematic representation of effects of irradiation on organic molecules

2. Molecular Scission: The molecule merely becomes noticeably shorter, perhaps 1/2 or 1/3 its original length. This reduces tensile and impact strengths, in extreme cases, to negligible values.

3. Polymerization: The passage of high-energy particles or photons through a long-chain polymer will apparently cause cross-linking between portions of the molecular chains when this is chemically feasible. This results in appreciable stiffening of flexible materials, hardening of greases, and solidification of fluids. The effects produced in many cases are similar to those observed under prolonged high-temperature heating.

Fig. 1 shows these processes schematically.

Techniques

It seems to be an unfortunate fact that many of the locations requiring electrically insulated apparatus differ widely in the composition of the radiations at that point. Some locations involve large amounts of gamma radiation with some fast neutrons; others, fast and thermal neutrons with fewer gammas, etc. While some work has been done in an attempt

Table II. Source Characteristics

Source	Particle	Energy Distribution	Optimum Sample Size	Approximate Dose Rate, r/hr	Normal Ambient, C	Changes in Ambient
General Electric Research Laboratory external electron beam	β	0.8-Mev peak sinusoid	not over 2-inch diameter, 0.040 inch thick (density)	$\approx 10^{9}$	20, air	possible
Oak Ridge reactor	γ, n	fission spectrum	length = 6 inches, $3^1/_2$ inches squared, also 2-inch diameter	$\approx 10^{6}$ (γ), 10^{12} neutrons per centimeter²-second	160 to 180, air	difficult
Brookhaven reactor	γ, β n	fission spectrum	length = 6 inches, $3^1/_2$ inches squared, also 2-inch diameter	$\approx 5\times10^{6}$ (γ), 5×10^{12} neutrons per centimeter²-second	160 to 180, air	difficult
Hanford discharged fuel elements	γ	fission product spectrum	length = 3 to 10 inches, diameter = 1 to 2 inches	1 to 2×10^{5}	20, water	practical
MTR discharged fuel elements	γ	fission product spectrum	length = 10 inches, 3 inches squared	8×10^{6}	20, water	practical
KAPL Cobalt 60 source	γ	1.33 Mev 1.17 Mev	length = 6 inches, diameter = $1^1/_4$ inches	2×10^{6}	20, water	practical
			annular volume	0.5×10^{6}		

Table III. High-Level Dosimeters

Type	Measures	Technique	Absolute	Precision, Per Cent	Useful Range
Ferrous-ferric	integrated dose	chemical analysis (colorimetric)	no (considered as standard)	± 2	to 2×10⁶ roentgens
Cerous-ceric	integrated dose	chemical analysis (colorimetric)	no	± 2	to 10⁷ roentgens
Ionization chamber	dose rate	current measurement	no	±10	to 5×10⁶ r/hr
Scintillation counter	dose rate	current measurement	no	±10	to 2×10⁶ r/hr
PVC film with methyl violet dye	integrated dose	light transmission (no chemical treatment)	no	±10	2×10⁵ to 2×10⁶ roentgens
Calorimeter	dose rate	power input	yes (subject to correction factors)	± 5	very high

to correlate the effects of fast neutrons with thermal neutrons, and thermal neutrons with gamma rays, results are not yet conclusive.[8] Frequently, the radiation levels estimated to occur at a given piece of equipment cannot readily be duplicated by any conveniently accessible radiation source. For a given mixture of neutrons and gamma rays, it is frequently necessary to test with too much of one or the other component of the flux. The following sources have been used in this laboratory; see Table II.

ELECTRON ACCELERATORS

Electron Accelerators provide the possibility of a very large dosage of electrons, of the order of 10^9 r/hr, for relatively thin samples. The effect of high-energy gamma rays is, to a first approximation, the same as that of high-energy electrons in a material. The use of electron accelerators thus offers the possibility of highly accelerated tests on thin samples of limited area.

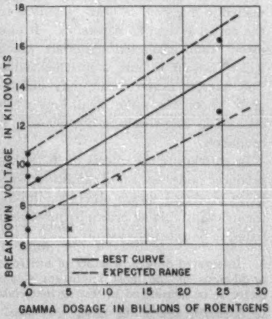


Fig. 2. Effect of gamma irradiation on short-time voltage breakdown in mica-glass-silicone structure. Circles represent data taken in absence of radiation field; crosses, data taken in radiation field of 2.4×10^6 r/hr. Insulation: three half-laps of 4-mil glass-backed mica tape, impregnated with General Electric SR-32 silicone resin

NUCLEAR REACTORS

Both of the Oak Ridge and Brookhaven reactors are air-cooled. It is necessary to construct an experiment which either fits the lattice tubes or one of the experimental holes. These are in no cases greater than a few inches on a side and, for maximum flux conditions, are usually the same size as the lattice tube, which may be of the order of 2 inches in diameter. Relatively high thermal neutron fluxes are available in central locations in these reactors (of the order of 10^{12} neutrons per square centimeter per second), inescapably accompanied by an ill-defined quantity of fast neutrons and a photon flux of 1-Mev gamma rays of about 10^6 r/hr. Somewhat higher flux levels may be obtained in the Materials Testing Reactor (MTR) under substantially the same conditions. Experiments at temperatures different from those of the reactor ambient is difficult without a water-cooled experimental hole. These supply neutron fluxes at 20 C. The normal Brookhaven or Oak Ridge reactor ambient is 160 to 180 C.

GAMMA-IRRADIATION SOURCES

The Knolls Atomic Power Laboratory has used extensively its own source of 3,000 curies of radioactive Cobalt 60 which produces a 1.33 and 1.17-Mev gamma. The maximum flux obtainable with this source is about 2×10^6 r/hr at the center and about 0.5×10^6 r/hr over an annular space 6 inches high with an inner radius of 1$^1/_2$ inches and with an outer radius of 6 inches. Higher pure gamma intensities can be obtained at the MTR site by placing specimens close to expended fuel elements which have been recently removed from the reactor. Levels over 10^7 r/hr can be obtained for short times. Both of the foregoing sources are under at least 10 feet of water to provide sufficiently low radiation levels for personnel in the vicinity. Samples to be exposed must, therefore, be placed in water-tight containers. It is possible, of course, to provide suitable internal heaters so that the experiment can be run at any desired temperature.

Materials which have been exposed solely to gamma radiation do not become radioactive and therefore can be examined with ordinary laboratory equipment in the usual way. Those which have been exposed to neutron radiation usually become radioactive. This necessitates health-physics precautions to determine tolerable levels and, in the case of destructive tests, such as abrasion tests, the debris from the test must be contained.

The advantages of a gamma source are several. The Cobalt 60 source, with a half-life of 5.2 years, requires calibration only at very infrequent intervals. Once a given pattern of active elements is positioned and measurements are made of the intensity at representative distances around the source, dose rates are established for an indefinite period. A further advantage is that the irradiation, generally speaking, is relatively inexpensive per roentgen exposure of the samples. This is primarily because the equipment requires almost no maintenance and no adjustments. There is nothing to wear out or to get out of order.

DOSIMETRY

Measurement of the dose rate and accumulated dosage during an experiment is obviously of great importance. Table III lists some of the more popular techniques. At this laboratory, the ferrous-

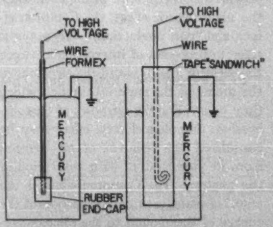


Fig. 3. Dielectric breakdown specimens in test position

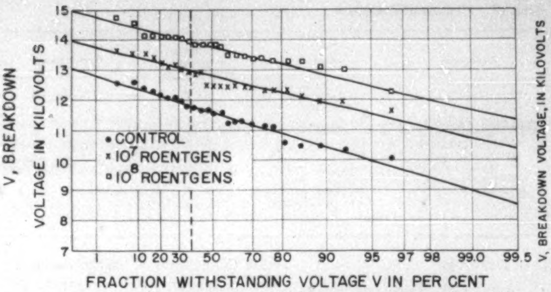


Fig. 4. Short-time dielectric breakdown of 9.5-mil polyvinyl chloride pressure-sensitive tape, showing effect of gamma irradiation

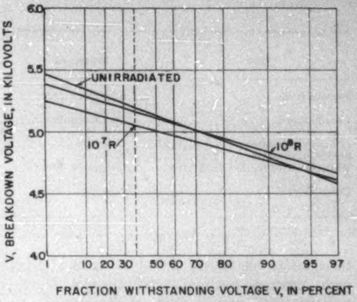


Fig. 6. Effect of gamma irradiation on dielectric breakdown of 5-mil cellulose acetate pressure-sensitive tape. Points omitted for clarity

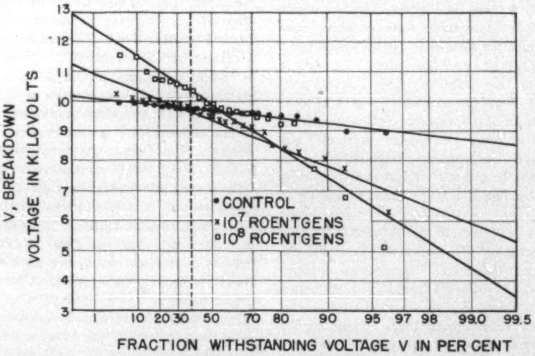


Fig. 5. Short-time dielectric breakdown of 6-mil polyethylene pressure-sensitive tape, showing effect of gamma irradiation

ferric dosimeter and the polyvinyl chloride film dosimeter are popular because of their relative convenience of use.[7]

There are three means whereby the measured dose delivered to a thickness of insulation adjacent to a metal surface by gamma rays may be increased above that observed in the interior of the insulation: the photoelectric effect, the Compton effect, and "pair production." These terms refer to processes of interaction between the incident photons and the electrons in the atoms of the material through which they pass. For the purposes of this discussion, the photoelectric effect may be considered the removal of an electron from its parent atom with absorption of the photon. The Compton effect is the result of transfer of only a portion of the energy of the photon to the electron, removing the electron from its parent atom and producing a new photon of lower energy than the incident gamma ray. Pair production occurs when annihilation of a photon leads to production of an electron and a positive electron, or positron. For a complete discussion of the nature of these processes, see reference 8. Of the three modes of electron production by incident gamma or X rays, only the photoelectric and Compton effects are of importance in insulation studies. Pair production has a threshold at 1 Mev and is negligible at photon energies below 5,000,000 electron volts. For the most part, the electrons which gain energy by any of the foregoing processes dissipate this energy within a few mils, or less, of the point of interaction.

Near the surface, some of the electrons are ejected with appreciable energies. The numbers and energies of electrons emitted by matter under exposure to gamma rays depend on the atomic number of the bombarded atoms (a measure of the electron density) and the energy of the incident photons. Prediction of the electron energy spectrum in any given situation is difficult because of the complex sequence of energy loss by the photon.[9] Suffice it to say that, for a metallic conductor in contact with an organic insulator, there are more electrons emitted by the conductor than by the insulator. The result is that the insulator receives a higher dosage in the region near the metal than it does in its interior. It should be remembered that the range of electrons in solids is sufficiently small to make this inhomogeneity appreciable, sometimes enhancing the dosage at the metal-insulator interface by as much as 100 per cent. The mechanism therefore makes it imperative that any dosimetry be conducted in a system as nearly identical to the actual system as is possible. It also means that insulation materials which will withstand, in bulk, a specified radiation flux for a given time may disintegrate under the same conditions when exposed in thin sheets in contact with a conductor.

In correlating the effects on materials produced under one set of conditions with those produced under other conditions, the following guiding principles have been found to be generally useful:

1. Integrated dosage, whether of neutrons or photons, is the most important single factor to be considered in testing materials for radiation stability.

2. Dose rates may generally be neglected, except when exceedingly high dose rates, i.e., 10^9 r/hr are encountered. For example, gas evolution at high dose rates may produce bubbles within an insulation structure which will rupture the material. At lower rates, evolution of gases may proceed sufficiently slowly to permit the gas to diffuse through the structure.

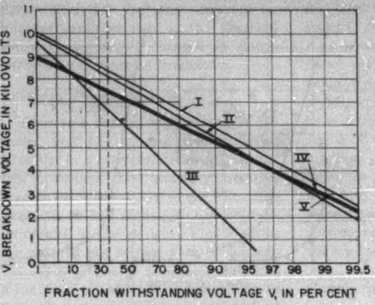


Fig. 7. Effect of gamma irradiation on dielectric breakdown of Formex-insulated copper wire. Points omitted for clarity

I—Annealed 30 minutes at 125 C
II—Annealed, then exposed to 10^7 roentgens
III—Annealed and exposed to 10^8 roentgens
IV—Annealed, exposed to 10^7 roentgens, then annealed again
V—Annealed, exposed to 10^8 roentgens, then annealed again

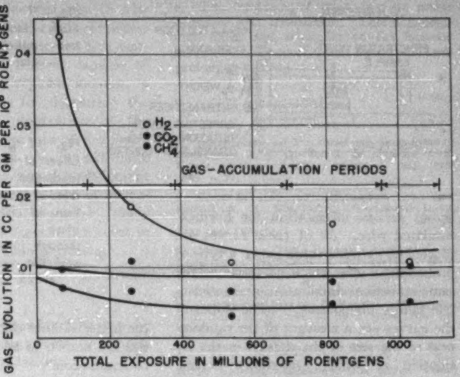


Fig. 8. Rate of gas evolution from General Electric SR-32 silicone resin as a function of total accumulated dosage

3. Radiation-induced processes result as a consequence of individual particle energies corresponding to temperatures of the order of thousands of C. It would therefore be expected that changes in the ambient temperature during a given irradiation would have little effect on the radiation-induced deterioration. On the other hand, it would be expected that elevated temperature would accelerate postirradiation reactions.

DIELECTRIC STRENGTH

The effect of gamma radiation on the dielectric strength of an insulation structure composed of mica-glass tape impregnated with silicone resin is shown in Fig. 2. The best curve was obtained by connecting the mean values of the data obtained with no irradiation and after an exposure to 24.5×10^8 roentgens. The expected range was obtained by assuming that the relative mean deviation obtained for the unirradiated specimens, 16 per cent, could be expected for all radiation dosages. It is interesting to note that the band encompassed by the relative-mean-deviation lines (shown dotted) includes all but two of the points obtained by breaking down the insulation in the absence of irradiation. Two data were obtained with the specimens in the radiation field. As could be expected, presence of radiation lowers the breakdown voltage as a result of ionization.

The increase in dielectric strength observed with the silicone-mica-glass insulation is consistent with the observation that the particular silicone employed undergoes additional polymerization under irradiation. The larger molecules which result from this process may be capable of withstanding larger voltages than the smaller polymer molecules which are obtained by ordinary curing procedures. This has been observed with a series of hydrocarbon liquids.[10] It should be noted, however, that the change in dielectric strength is only about 25 volts per 100,000,000 roentgens, far smaller than that usually observed in irradiation of pure organic insulation structures.

Because of the considerable interest in organic insulation, a series of a-c breakdown tests was made on polyethylene, polyvinyl chloride (PVC) and cellulose acetate films, and on Formex-insulated wire. Film specimens were prepared by making a "sandwich" of copper wire between two thicknesses of standard pressure-sensitive tape, as shown in Fig. 3. The sandwich was then immersed in mercury and voltage applied to the copper wire. The samples of Formex-insulated wire were cut, stripped of insulation at one end, and annealed at 125 C for 30 minutes. A portion of a rubber eraser was placed over the other end of the wire, and served as an insulator during immersion in the mercury electrode of the breakdown tester. Voltage rise was controlled at 0.5 kv per second. Breakdown normally took place along the length of the wire, it was observed.

Figs. 4, 5, and 6 display data obtained after exposure of films of polyvinyl chloride, polyethylene, and cellulose acetate to 10^7 and 10^8 roentgens. Fig. 7

Table IV. Effect of Gamma Radiation on Power Factor of Pump Winding Facsimile; Measurements at 1,000 Cycles per Second

Dosage, Roentgens $\times10^{-8}$	Power Factor, Per Cent
0	1.77
1.5	1.24
5.0	1.24

Table V. Gas Evolution From Thin Films of SR-32 During Pyrolysis Film Thickness to 10^{-3} Millimeters

Time	Temperature, C	Evolution Rate, Microliters per Minute	Composition, Per Cent
15 minutes	90	0.5	Methyl trimer, tetramethylsilane and low molecular weight silanes
30 minutes	95		
40 minutes	135	0.25	95 C_2H_4, 4 H_2O, 1 trimer
70 hours	135	0.0042	95 C_2H_4, 4 trimer
3 hours	150		C_2H_4, H_2O, cyclic trimer, tetramers
16 hours	220		50 C_2H_4, 20 hexamethyltricyclosilane

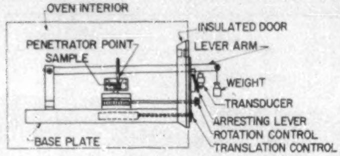


Fig. 9 (left). High-temperature penetrometer

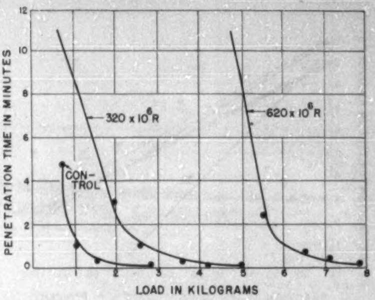


Fig. 11 (right). Effect of gamma irradiation on rate of penetration of 90-degree point through typical asphaltic varnish at 25 C

shows similar information for Formex-insulated wire. All of these curves are plotted on extreme-value paper, a type of probability paper which conforms to the same statistical distribution as do dielectric failure phenomena.[11] The slopes of the curves are a measure of the randomness of the size of the defects in the insulation at which breakdown occurs. An increase in the absolute value of the slope represents an increase in the range of size of defects. The intersections of the curves with the dotted vertical line are the most probable values of breakdown voltage and indicate the most probable size of the defects.

Irradiation of polyvinyl chloride, Fig. 4, seems to have little effect on the slopes of the curves; hence the distribution of the size of the defects over the film is unaffected. The size of the defects appears to decrease with irradiation as the breakdown voltage increases.

Pressure-sensitive polyethylene tapes show no definite change in the most-probable breakdown voltage with radiation dosage; see Fig. 5. However, there is a perceptible increase in the range of size of defects, as reflected by the increase in the slopes of the curves.

In the case of cellulose acetate, Fig. 6, it is interesting to note that no conclusive effect on the breakdown voltage can be seen, even with the highly magnified ordinate. For improved readability, the points have been omitted from the curves. It is somewhat surprising to observe so little change in the dielectric properties of the material since the mechanical properties are known to be extremely sensitive to irradiation.[4]

Fig. 7 compares the effects of postirradiation annealing on Formex-insulated wire with the effects of irradiation alone. From curves I, II, and III, it may be concluded that irradiation not only lowers the most probable breakdown voltage but also increases the range of sizes of the defects in the wire coating. However, upon annealing the irradiated wire samples for 30 minutes at 125 C, a surprising change takes place. The most probable breakdown voltage tends toward a constant value of about 7.5 kv, and the range of sizes of defects is diminished to a level which is slightly lower than that in the unirradiated material; see curves IV and V. A possible hypothesis which may explain this behavior is that the defects introduced by irradiation and those present after the initial annealing process are somehow combined and removed from the enamel by the action of the postirradiation annealing. A detailed study of this phenomenon has been made.

It may be concluded that, from a technical standpoint, the changes in dielectric breakdown behavior of a particular insulating material under irradiation are difficult to predict. As might be expected from the very different chemical compositions of the materials tested, each shows a different behavior after irradiation. Whether chemical processes have been initiated which will cause progressive deterioration and consequent lower breakdown voltages after aging for longer times cannot now be stated. It is evident, however, that further investigation of each specific compound will be necessary to determine completely its properties after irradiation.

Power Factor

Power-factor determinations have been used as an index of insulation quality for many years. Under certain circumstances, valuable conclusions can be drawn from their use. In other cases, interpretations of the measurements are inconclusive. The work of Fuoss[12] has shown a correlation between measurements of this kind and molecular structure.

Measurements in this investigation have been taken both on composite insulation structures and on thin films of varnish resin alone. Table IV shows the power-factor changes which have occurred in a facsimile of an electromagnetic pump winding in which the insulation was mica-and-glass tape impregnated with silicone resin. This was a bar-and-slot type of construction, essentially the same as is used in the medium-sized induction motor. It is to be noted that the changes are similar in many respects to those which would be observed by the thermal aging of the resins.

To date no consistent pattern has been observed relating observed power-factor changes to irradiation dosage, even though careful attention has been paid to the control of experimental variables such as humidity, temperature, etc. Frequencies from a few hundred to 50 megacycles have been used. It has been con-

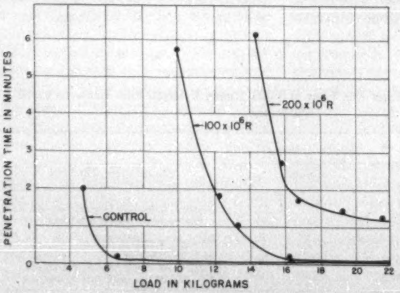


Fig. 10. Effect of gamma irradiation on rate of penetration of 1/16-inch steel ball through typical phenolic varnish at 70 C

Vol,
Pt. 1
1955

cluded that, for the more common insulating varnishes used in the electrical industry, significant changes in power factor do not occur on irradiation.

PHYSICAL DETERIORATION

In certain design applications, it becomes desirable to seal the insulating structure in a hermetic enclosure. Under this condition, the tendency of solid organic materials to evolve gas when subjected to nuclear radiations can be a determinant in the design. Fig. 8 shows the rate of gas release from a silicone-varnish-impregnated facsimile winding as a function of time. It is interesting to note that the quantity of gas evolved shows no diminution in rate after a period of about 1 month and that the fractional composition seems then to be approximately constant. High-temperature pyrolysis of the material gives a markedly different grouping of molecular fractions than is observed under nuclear radiation; see Table V. It would therefore be concluded that the processes induced by radiation are substantially different in kind than normal deterioration of insulating materials caused by temperature.

HARDNESS

As has previously been pointed out, mechanical changes of various kinds can occur depending upon the nature of the material. Sisman and Bopp have tabulated a considerable number of mechanical property changes for some of the commoner organic plastics.[5] The approach in this investigation has been somewhat different. It was felt that it would be desirable to determine the loss in viscosity in some typical electrical varnishes. A penetrometer was devised which consisted essentially of a lever capable of being weighted by a given amount, a stylus, and a means for measuring the time required for penetration of the stylus of a given thickness of irradiated material. Fig. 9 shows, in schematic form, the construction of this equipment and data taken with this type of instrument are shown in Figs. 10 and 11. These indicate the great changes which occur in the hardness of the resin in question.

Conclusions

The following generalizations can be made concerning the effects of irradiation on electrical insulation:

1. For dosages of radiation up to 10^8 roentgens, polyethylene tape and Formex wire enamel undergo discernible decreases in their resistance to short-time voltage breakdown. Cellulose acetate shows little change under the same conditions. Polyvinyl chloride shows some distinctly favorable alterations after irradiation.

2. When levels of radiation exposure up to 10^{11} roentgens are to be expected, mica-and-glass tape impregnated with silicone resin—essentially an inorganic insulation structure—has been demonstrated to be virtually unaffected or slightly improved by the irradiation.

3. Power factor has not been found to be a particularly sensitive index to radiation-induced changes on the materials studied.

4. The evolution of gas from organic material varies with the material and can be copious. It is usually hydrogen, carbon dioxide, or methane. Care must be observed in the design of sealed equipment placed in radiation fields.

5. The hardness of most organic resins increases with irradiation. This is a common effect and is similar to the increase in viscosity observed when there has been prolonged heating.

References

1. Irradiation of Polymers by High-Energy Electrons, E. J. Lawton, A. M. Bueche, J. S. Balwit. *Nature*, London, England, vol. 172, 1953, p. 76.

2. Cross-Linking of Polythene by Pile Radiation, A. Charlesby. *Proceedings*, Royal Society of London, England, vol. 215, pt. A, 1952, pp. 187–214.

3. Effect of High-Energy Radiation on Long-Chain Polymers, A. Charlesby. *Nature*, London, England, vol. 171, 1953, p. 167.

4. Testing Electrical Insulation for Use in Gamma-Ray Fields, C. Mannal. *Nucleonics*, New York, N. Y., vol. 12, no. 6, June 1954, pp. 49–53.

5. Physical Properties of Irradiated Plastics, O. Sisman, C. D. Bopp. *ORNL-928*, Oak Ridge National Laboratory, Oak Ridge, Tenn., 1951.

6. Radiation Effects on Reactor Materials. Non-Metals, V. P. Calkins. *Nucleonics*, New York, N. Y., vol. 12, no. 9, Sept. 1954. pp. 9–15.

7. Gamma-Ray Dosimetry with Polyvinyl Chloride Films, E. J. Henley, A. A. Miller. *Ibid.*, vol. 9, no. 12, Dec. 1951, pp. 62–65.

8. Radioactivity and Nuclear Physics, J. M. Cork. D. Van Nostrand Company, Inc., New York, N. Y., 1947, pp. 142–55.

9. Electron Energy Distributions Produced by Gamma-Rays, H. E. Johns, J. E. Till, D. V. Cormack. *Nucleonics*, New York, N. Y., vol. 12, no. 10, Oct. 1954, pp. 40–46.

10. Electric Strengths and Molecular Structure of Saturated Liquid Hydrocarbons, R. W. Crowe, A. H. Sharbaugh, R. V. Bragg. *Journal of Applied Physics*, New York, N. Y., vol. 25, 1954, pp. 1480–82.

11. Statistical Theory of Extreme Values and Some Practical Applications, E. J. Gumbel. Applied Mathematics Series no. *33*, U. S. National Bureau of Standards, Washington, D. C., 1954, p. 40.

12. Series on Electrical Properties of Solids, R. M. Fuoss. *Journal*, American Chemical Society, New York, N. Y., vol. 59, 1937, pp. 1703–09.

Transfer Functions for Amplitude-Modulated Signals

H. HELLERMAN
ASSOCIATE MEMBER AIEE

THE purpose of this paper is to derive the general theoretical requirements which must be met by linear networks which are to perform specified operations on amplitude modulated signals. In particular, an exact correspondence will be established between a realizable transfer function $A(s)$ which performs a given operation on any modulating function $f_m(t)$ and the network which performs an identical operation on the envelope of the modulated function $f_m(t)\sin\omega_0 t$ where $\omega_0/2\pi$ is the carrier frequency.

Consider the general configuration shown in Fig. 1. The relation between output and input of the system is given by the familiar transform relation

$$F_0(s)=F_i(s)A(s) \tag{1}$$

A straightforward procedure to find the corresponding network function for a modulated input of the form $f_i(t)\sin\omega_0 t$ is to equate the output transform to the product of the desired transfer function $B(s)$ and the transform of the input. This procedure leads to the following expression:

$$B(s)=\frac{F_0(s-j\omega_0)-F_0(s+j\omega_0)}{F_i(s-j\omega_0)-F_i(s+j\omega_0)} \tag{2}$$

With the use of equation 1 for F_0, equation 2 may be written

---

Paper 55-706, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted June 20, 1955; made available for printing August 18, 1955.

H. HELLERMAN is with Syracuse University, Syracuse, N. Y.

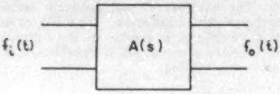


Fig. 1. Four-terminal network

$$B(s) = \frac{F_i(s-j\omega_0)A(s-j\omega_0)-F_i(s+j\omega_0)A(s+j\omega_0)}{F_i(s-j\omega_0)-F_i(s+j\omega_0)} \tag{3}$$

It is required that $B(s)$, the desired transfer function, should be independent of F_i. This can be true only if $A(s-j\omega_0) = A(s+j\omega_0)$ and then $B(s) = A(s-j\omega_0)$. Now if $A(s)$ is realizable its poles and zeros occur in conjugate pairs. The function $A(s-j\omega_0)$ cannot then have its poles occur in conjugate pairs and hence $B(s)$ is not realizable. (Actually this proof breaks down in the degenerate case where $A(s)$ is independent of s.)

Since the straightforward procedure was not successful, it is natural to seek another configuration besides the general 2-terminal pair network which may perform the desired operation. Such a general configuration is shown in Fig. 2. Here it is seen that the modulated function appears in two forms. The second form $f_m(t)\cos\omega_0 t$ will be called the quadrature function. Its transform is given by

$$L[f_m(t)\cos\omega_0 t] = \frac{F_m(s-j\omega_0)+F_m(s+j\omega_0)}{2} \tag{4}$$

The modulated signal and its quadrature function are separately applied to networks whose transfer functions are $A_1(s)$ and $A_2(s)$. Finally the outputs of these networks are added. If the networks have transfer functions satisfying the following relations

$$A_1(s) = \frac{A(s-j\omega_0)+A(s+j\omega_0)}{2}$$
$$A_2(s) = \frac{A(s-j\omega_0)-A(s+j\omega_0)}{2j} \tag{5}$$

then the final output will be $f_0(t)\sin\omega_0 t$.

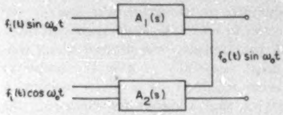


Fig. 2. Configuration for producing A(s) operation on the envelope of a modulated signal

If the networks $A_1(s)$ and $A_2(s)$ are realizable, then operations corresponding to $A(s)$ on the modulating function may be performed on the modulated function. The realizability of the two network functions of equation 5 may be established by noting first that the poles of $A(s-j\omega_0)$ and $A(s+j\omega_0)$ together constitute the poles of $A_1(s)$ and $A_2(s)$. Both of these functions have their poles in the left half-plane since they are simply the poles of $A(s)$ displaced parallel to the imaginary s axis. Fig. 3 shows the pole locations of $A(s)$, $A(s-j\omega_0)$, $A(s+j\omega_0)$ and $A_1(s)$ for a typical pair of poles of $A(s)$. $A_1(s)$ is seen to have poles which occur in complex conjugate pairs and lie in the left-half s-plane and thus satisfy the realizability criterion. A further point of interest is that both $A_1(s)$ and $A_2(s)$ have the same poles which also means they have the same network determinant to within a power of s.

To complete the proof of the realizability of the functions $A_1(s)$ and $A_2(s)$ it is necessary to show that the zeros of these functions occur in complex conjugate pairs. This is equivalent to showing that the coefficients of the numerator polynomial of $A_1(s)$ are all real and the coefficients of the numerator polynomial of $A_2(s)$ are all imaginary. These requirements are met if $A(s)$ is realizable since a typical term of the numerator polynomial of $A_1(s)$ is

$$a_k(s-j\omega_0)^k b_n(s+j\omega_0)^n + a_k(s+j\omega_0)^k b_n(s-j\omega_0)^n \tag{6}$$

Equation 6 is a polynomial in s with real coefficients and since the numerator of $A_1(s)$ consists of the sum of such polynomials, this numerator satisfies the conditions necessary to make $A_1(s)$ realizable. A similar line of reasoning will establish the realizability of $A_2(s)$.

If the output is desired in the form $f_0(t)\cos\omega_0 t$ instead of $f_0(t)\sin\omega_0 t$ it is merely necessary to interchange the inputs to the A_1 and A_2 networks. Thus, by the use of two sets of networks both the desired modulated output and its quadrature function may be obtained. This pair is suitable for feeding another network, say for the purpose of producing cascaded operations.

These results are concerned with performing exact specified operations on amplitude-modulated signals. This in general requires the quadrature input signal. However, if only approximate results are required as is often true in communications work, then the operation may be performed without a quadrature signal. This case has been explored by Agrain, Teare, and Williams.[1] Their work indicates that when the modulating function has a spectrum which is principally confined to a band which is a small fraction of the carrier frequency, the operation corresponding to $A(s)$ may be performed without a quadrature signal.

Example

Find the networks $A_1(s)$ and $A_2(s)$ which will realize the transfer function corresponding to $A(s) = 1/T(s+1/T)$

From equation 5

$$A_1(s) = \frac{\left(s+\frac{1}{T}\right)}{T\left[\left(s+\frac{1}{T}\right)^2+\omega_0\right]}$$

$$A_2(s) = \frac{\omega_0}{T\left[\left(s+\frac{1}{T}\right)^2+\omega_0^2\right]}$$

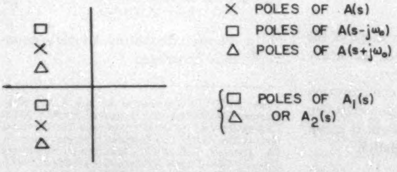


Fig. 3. Pole locations for a typical pair of A(s) poles

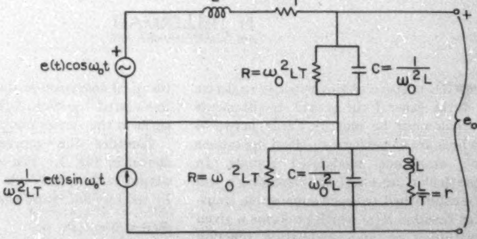


Fig. 4. A configuration corresponding to A(s)=1/1+sT

The poles of these functions suggest a uniformly dissipative parallel inductance-capacitance network. Because the A_1 and A_2 functions have the same poles it is expected that the same type of network could be used for both, with differences being the terminals and quantities which are specified as input and output. Such is indeed the case. One possible circuit realizing the required responses is shown in Fig. 4.

Conclusions

The foregoing theory showed the conditions which must be met to realize exactly a transform operation on the envelope of a signal of the form $f_m(t)$ sin $\omega_0 t$ corresponding to the transfer function $A(s)$ for the modulating function. It was shown that a quadrature signal of the form $f_m(t)$ cos $\omega_0 t$ was essential to perform the desired operation exactly. Assuming the quadrature function is available, two transfer functions which were shown to be realizable are necessary. These may be calculated in a straightforward manner from $A(s)$.

Reference

1. Generalized Theory of the Band-Pass Low Pass Analogy, **Agrain, Teare, Williams.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., Oct. 1949.

---

ol. 74
. 1
55

A Polar Relay Using Momentum Transfer

H. L. GARBARINO
ASSOCIATE MEMBER AIEE

K. E. BISSHOPP
NONMEMBER AIEE

A PRINCIPAL application of magnetic polar relays is to circuits where high operating speeds are required. One common approach for obtaining high-speed, bounce-free operation of relay contacts has utilized reduction of the armature size and mass. Disadvantages in reducing armature size are the resulting reduction of relay contact forces and the need for smaller mechanical tolerances.

The application of the momentum-transfer principle to a polar relay is described, the impact behavior is analyzed and tests on simple experimental relays are summarized. These results show promise for improved relay performance in some respects, such as the high contact forces which are readily obtainable and the freedom from contact bounce. The momentum-transfer principle may have value in other switching devices such as circuit breakers.

Application of Momentum Transfer

The principle of momentum transfer is exemplified by three collinear balls of equal mass and perfect elasticity, two of which, the second and third, are adjacent and stationary. If the first ball is moving at some initial velocity and strikes the second, it will stop, and the third ball will move with the original velocity of the first ball. Ideally, the ball in the middle of the three would not move. In a polar relay, the three balls become contacts. The transfer of momentum permits fast closing of one pair of contacts and almost simultaneous opening of another pair, as shown by tests of rather crude models. The pair of contacts which are being closed should not bounce or chatter, a condition which could be obtained only if the masses are equal and perfectly elastic during impact. Since no materials are perfectly elastic, some chatter may be expected. However, it has been found that bounce can still be eliminated even with nonideally elastic materials by making the middle mass slightly larger than the other two, according to a ratio which depends on the energy lost during impact.

Fig. 1 shows the manner in which an electromagnetic relay may be equipped with three equal masses as the contact elements. An insulated hammer at the end of the relay armature opens the circuit between the middle mass and one outer mass or contact and permits the second outer contact to remain against the middle one. When the armature is reversed, the open outer contact is released, and it flies in, striking the center mass. The second outer mass then breaks contact before being struck by the hammer. This contact is caught by the hammer when it attempts to return to the center contact under the force of its mounting spring. The three contacts are suspended from springs mounted in a stationary frame. Springs holding the outer masses must provide the stationary contact force between each outer contact and the center contact. By making the center spring comparatively stiff, deflection of the middle mass may be kept small even when only one outer contact is closed. Spring forces do not affect transfer of momentum because the impact forces are much higher.

In the use of the momentum-transfer principle, it is intended that all three contacts will be closed for a very short part of the period when the armature is changing its position. Although the armature of Fig. 1 is operated magnetically, this could be accomplished mechanically with similar contact performance.

Impact Relations

The series of impacts occurring during one contact operation are of interest. Impact equations are basically relationships among momentums. Behavior of the contacts between each series of impacts can be best described by using differential equations of forces. During these periods it is necessary to include spring forces.

For the impact period, consider three colinear masses as in Fig. 2, of which m_1 is moving with initial velocity V_1, and it is assumed that m_0 and m_2 have become stationary since the previous operation. The system can be analyzed by computing the momentum transfer during each collision. Study of Fig. 2 shows that the first impact occurs between m_1 and m_0. Also it is assumed that impact forces between m_1 and m_0 have reached a maximum and decreased to zero before forces are built up between m_0 and m_2.

General relations among any initial and final velocities of two masses, as given by Osgood,[1] are

$$V_1' = \frac{(m_1 - em_0)V_1 + (1+e)m_0V_0}{m_0 + m_1} \tag{1}$$

$$V_0'' = \frac{(1+e)m_1V_1 + (m_0 - em_1)V_0}{m_0 + m_1} \tag{2}$$

The quantity V_0'' is the velocity of mass m_0 after being struck by m_1 but before impact forces are built up between m_0

Paper 55-175, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 31–February 4, 1955. Manuscript submitted October 14, 1954; made available for printing March 1, 1955.

H. L. Garbarino and K. E. Bisshopp are with the Armour Research Foundation of Illinois Institute of Technology, Chicago, Ill.

The investigation upon which this paper is based was sponsored by the Teletype Corporation. The encouragement and guidance of the Teletype staff members are greatly appreciated. Acknowledgment is also made of the valuable assistance given by A. R. Vogt of the Foundation.

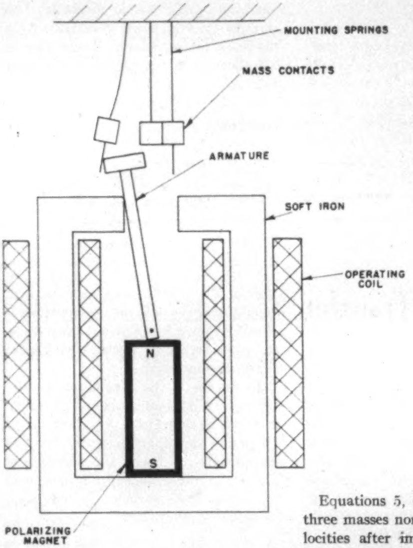


Fig. 1 (left). Momentum transfer in a polar relay

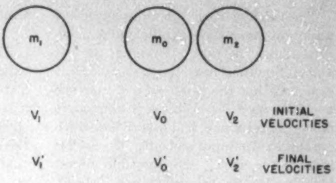


Fig. 2 (right). Momentum transfer of masses

and m_2. The quantity e is the coefficient of restitution. The simplest definition for this quantity is obtained for the case when a struck mass such as m_0 is infinite. Then e is the ratio of final to initial velocities of m_1. More generally, e is the ratio of the integrated force-time product as impact force decreases from its maximum value to zero, to the integrated force-time product as impact force increases from zero to its maximum value.

When mass m_0 with velocity V_0'' collides with mass m_2, the general relations for any initial and final velocities give

$$V_0'=\frac{(m_0-em_2)V_0''+(1+e)m_2V_2}{m_0+m_2} \tag{3}$$

$$V_2'=\frac{(1+e)m_0V_0''+(m_2-em_0)V_2}{m_0+m_2} \tag{4}$$

Since the initial conditions for the particular case of interest are $V_1 \neq 0$, and $V_0 = V_2 = 0$, the velocities after m_2 is struck are

$$V_1'=\frac{m_1-em_0}{m_0+m_1}V_1 \tag{5}$$

V_0' is found by substituting for V_0'' in equation 3 from equation 2.

$$V_0'=\frac{(1+e)m_1(m_0-em_2)}{(m_0+m_1)(m_0+m_2)}V_1 \tag{6}$$

$$V_2'=\frac{(1+e)^2m_0m_1}{(m_0+m_1)(m_0+m_2)}V_1 \tag{7}$$

Equations 5, 6, and 7 show that the three masses normally have different velocities after impact. Each mass then acts independently until a new sequence of impacts occurs. The equations show that ideal performance for equal masses is not obtained because the coefficient of restitution e is always less than unity. This means that V_1' and V_2' are not zero after m_2 has been thrown outward. Also, since V_1' and V_0' are unequal for equal masses, it will be found that another impact will occur between m_0 and m_1, followed by separation, so that V_1' and V_0' are not the final velocities.

The impact problem was studied to find if there might be ratios of masses for which bouncing between m_1 and m_0 can be eliminated after m_2 has been thrown off. This would be obtained if V_1' and V_0' were equal. Since bouncing must be prevented during operation in both directions, it is reasonable to require that the outer masses be equal: $m_1=m_2$ for symmetry. Equations 5, 6, and 7 then become

$$V_1'=\frac{m_1-em_0}{m_0+m_1}V_1 \tag{8}$$

$$V_0'=\frac{(1+e)m_1(m_0-em_1)}{(m_0+m_1)^2}V_1 \tag{9}$$

$$V_2'=\frac{(1+e)^2m_0m_1}{(m_0+m_1)^2}V_1 \tag{10}$$

It is evident that no values for m_1 and m_0 can make V_1' and V_0' both equal zero if e is not unity. However, there is a condition for which $V_1'=V_0'$. Equating 8 and 9 give

$$\frac{m_1-em_0}{m_0+m_1}=\frac{(1+e)m_1(m_0-em_1)}{(m_0+m_1)^2} \tag{11}$$

If there is a solution, both sides of equation 11 must be either positive or negative. Simplifying equation 11 and setting

$$x=\frac{m_0}{m_1} \tag{12}$$

gives the quadratic equation in x

$$x^2+2x-\frac{1+e+e^2}{e}=0 \tag{13}$$

By the quadratic equation, the roots of equation 13 are

$$x=-1\pm\sqrt{1+\frac{1+e+e^2}{e}} \tag{14}$$

The existence of a useful solution to the problem requires that these roots be real, and that at least one root be positive because a negative ratio of masses is meaningless. Since the absolute value of the radical term is greater than unity, there is one positive and one negative value for x. Using the positive sign and simplifying gives

$$x=-1+\frac{1+e}{\sqrt{e}} \tag{15}$$

Table I which gives values of the ideal mass ratio as a function of the coefficient of restitution was obtained from equation 15. For prevention of bounce, m_0 should be somewhat greater than $m_1=m_2$. The results check equation 11 in that m_1 is greater than em_0 and m_0 is greater than em_1 so that both sides of equation 11 are positive. As another check note that the mass ratio is unity when e is unity.

Table I. Coefficient of Restitution and Ideal Mass Ratios

e	$x=m_0/m_1$
1.0	1.0
0.9	1.005
0.8	1.01
0.7	1.03
0.6	1.07
0.5	1.12
0.4	1.21
0.3	1.37
0.2	1.68
0.1	2.48
0	Infinite

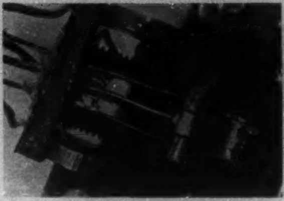


Fig. 3. Test model

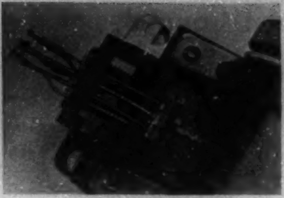


Fig. 4. Test model

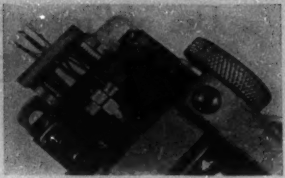


Fig. 5. Test model

To find the velocities V_1', V_0', and V_2' in terms of e and x, equations 8, 9, and 10 are readily changed to

$$V_1' = \frac{1-ex}{1+x} V_1 \tag{16}$$

$$V_0' = \frac{(1+e)(x-e)}{(1+x)^2} V_1 \tag{17}$$

$$V_2' = \frac{(1+e)x}{(1+x)^2} V_1 \tag{18}$$

Substituting for x from equation 15 and simplifying gives final velocities as a function of e when the ideal mass ratios are used

$$V_1' = V_0' = \sqrt{e}(1-\sqrt{e})V_1 \tag{19}$$

$$V_2' = \sqrt{e}(1+e-\sqrt{e})V_1 \tag{20}$$

An experimental method for finding coefficient of restitution e is to vary the mass ratio m_0/m_1 and to find that value at which no bounce occurs. The coefficient can then be determined from Table I. It is expected that e depends on contact mounts and materials, and is probably affected by the presence of rivets.

Simple examples show the theoretical behavior of the contact system when a mass ratio above or below the ideal value is used. An example of a ratio too low is the case of all masses equal and e less than unity. For this case, $x = m_0/m_1 = 1$. Equations 16 and 17 give

$$V_1' = \frac{1-e}{2} V_1$$

$$V_0' = \frac{(1+e)(1-e)}{4} V_1 \tag{21}$$

These conditions hold immediately after m_2 has been thrown outward. Both V_1' and V_0' are positive (i.e., in the same direction as the original V_1), but V_1' is greater than V_0'. This means that another compression between m_1 and m_0 will occur, and they will then bounce apart.

An example of a mass ratio higher than the ideal value is that of x greater than unity and e equal to unity. Then equations 16 and 17 become

$$V_1' = -\frac{x-1}{x+1} V_1$$

$$V_0' = \frac{2(x-1)^2}{(x+1)} V_1 \tag{22}$$

These show that V_1' is negative and V_0' is positive so that the masses m_1 and m_0 move apart after the initial impact between m_1 with velocity V_1, and m_0 with velocity zero. Eventually spring forces acting on each of the masses m_1 and m_0 will bring them together for another impact.

Model Tests

Several tests have been made on rather crude experimental models which use the momentum-transfer principle. To evaluate mechanical performance in the simplest manner, special contacts were mounted in the magnetic structure of a Western Electric *255*-type polar telegraph relay. Among the many possible variations of contact configurations, those shown in Figs. 3, 4, and 5 were built and tested.

Contacts supported by piano-wire springs were used in all cases. The springs are fastened into holes drilled into the contact pieces. As the figures show, the wire supporting the middle spring is much heavier than those supporting the outer two. The outer wire springs are stressed so that a force of about 40 to 45 grams is applied between the center and each outer contact. The tops of the springs are brought through holes in a piece of bakelite and cemented. The electric leads are attached to the ends of the protruding spring wires. In every test model the armature, with or without a hammer, is insulated from the contacts. The stationary magnetic poles which act as armature stops had been set for about 7-mil travel. The free play between the armature and the two outer contacts was about 3 mils.

The models were tested with square wave pulses of current over a range of values suitable to the power requirements of the particular relay used. The models gave very short times between the first and last contact action, 0.2 to 0.4 millisecond. However the operating time from the change in signal current until the action of the contacts is much greater, 1.3 to 2.6 millisecond. Of the three models, that shown in Fig. 5 gave the best performance and was found to be the easiest to adjust. All models showed no contact bounce for intermediate signal currents, and little or no bounce for relatively high signal currents. This is rather remarkable in view of the high static contact forces to which the mounting springs were stressed.

Conclusions

1. The application of the momentum-transfer principle to the polar relay should yield little or no contact rebound. The time duration of contact action is very short, but operating times are considerably longer.
2. It is found that there exist ideal ratios of the masses for which contact rebound does not occur, even with materials which are not ideally elastic or with assemblies having appreciable impact losses.
3. Operation of rather crude experimental models has been in accordance with the results of the analysis.
4. The principle of momentum transfer involving three masses should have other applications in equipment where impact occurs and where it is desirable to reduce contact rebound, to obtain high contact forces, and to achieve good operating times.

Reference

1. MECHANICS, W. F. Osgood. Macmillan Company, New York, N. Y., 1937.

Use of Proportional Counters in X-Ray Diffraction

H. R. LAIRD
NONMEMBER AIEE

M. J. ZUNICK
NONMEMBER AIEE

IN X-RAY diffraction, the data are obtained by measuring the intensities diffracted at any given angle by a sample. In Fig. 1 is shown a typical experimental arrangement in which the beam from the X-ray tube target is diffracted at an angle 2θ from its original direction into the measuring device. This device can be moved from 0 to 165 degrees 2θ by a suitable mechanism as shown in Fig. 2, illustrating a present-day diffraction unit. Then, if the intensity is plotted on a chart as a function of the angle 2θ, a typical diffraction pattern is obtained as shown in Fig. 3. The information directly obtained from this chart concerns the spacing of the atomic planes within the tiny individual crystals present in the sample and the relative (uncorrected) intensities of diffraction associated with those planes. These spacings and intensities can be interpreted to give detailed information about the crystalline characteristics of the sample.

The intensity of these diffraction peaks is often extremely low, and the most commonly used kind of electronic measuring device has been the Geiger counter because of its great sensitivity and simplicity of use, although three different kinds of gaseous counters have been used: the ionization chamber, the proportional counter, and the Geiger counter.

Several authors, notably Lang and Knowles[1–3] have reported on results with the proportional counter. They used a cylindrical tube, as shown in Fig. 4, with the beam passing through the tube perpendicularly to the anode wire. With this construction, the electric field is better distributed than in the case where the beam enters parallel to the anode wire, as shown in Fig. 5. The result is that the pulse size distribution put out by the tube is much better with the side window tube, and this produces a much better plateau.

The three kinds of gaseous counter tubes just mentioned are defined in terms of their operating conditions, and the controlling factor is the voltage applied to the tube. The operation of these counter tubes will now be described.

When an X-ray quantum is absorbed by the gas in the counter, a certain number of ion pairs is formed. This number of ion pairs is proportional to the quantum energy and therefore inversely proportional to the wave length of the X-rays. If the voltage applied to the tube is just sufficient to pull all the ions to the electrodes before recombination, the tube is operating as an ionization chamber.

If, now, the voltage is raised until ionization by collision occurs, the original number of ions is multiplied by a factor called the gas amplification. This factor is a constant for a given applied voltage, and increases exponentially with voltage. It may be as large as 10,000. Over the range of voltages being discussed, the size of the output pulse maintains its proportionality to the energy of the original X-ray quantum. Hence, over this range of operation the tube is called a proportional counter.

As the voltage is raised still further, the discharge is no longer localized, but spreads throughout the whole length of the tube. The pulse size is now fixed, and is entirely independent of the original X-ray quantum. A short wave length radiation of high energy, or a long wave length radiation of low energy will both produce the same sized pulse. Under these conditions the tube is operating as a Geiger counter.

Experimental Work and Results

CONSTRUCTIONAL PROCEDURES

In the development work on counter tubes, every effort has been made to exclude foreign agencies in their construction and processing. Borosilicate glasses and metals from the ferrous and refractory groups were used in the construction almost entirely. Only copper and silver brazing materials were used. The use of these materials and of beryllium metal for windows allowed high temperature bake-out (450 degrees) during exhaust and prior to filling. All brazing operations were performed in vacuum, thus preventing oxide formation and resulting in clean metal surfaces. High purity gases were used and dried through liquid air traps which were a part of the exhaust and filling apparatus. The resultant permanently sealed tubes were found to give very consistent, long-lived operation, whether filled and operated as Geiger tubes or proportional counters.

CHOICE OF TUBE GEOMETRY

To produce pulses of uniform size for a given wave length of monochromatic radiation, it is necessary that all ions be accelerated the same amount, irrespective of the point where the particular ion is formed. This condition is closely approached with the side window tube shown in Fig. 4. In the case of the end window tube in Fig. 5 it can be shown that this condition is very far from fulfillment. Both Lang and Knowles considered the side window tube as the only style worthy of consideration.[1–3] The General Electric Company has tried the end window tube and has found the plateaus to have slopes as least 10 times as great as in the case of the side window tube.

For this reason attention has been confined to the side window tube, and to a modification of this style, the rectangular tube, shown in Fig. 6. In this counter the beam passes between two parallel plates, at right angles to the five anode wires. From the theoretical standpoint, the pulse sizes from this tube would not be expected to be quite as uniform as from the cylindrical side window tube. Experimental data on this point is at present meager, but a noticeable difference in pulse uniformity in the direction expected is indicated.

Considerable work has been done on the rectangular tube because it possesses one important advantage over the cylindrical side window tube, the greater length of gas path, which is 4 inches for

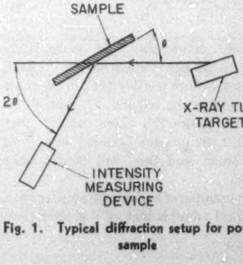


Fig. 1. Typical diffraction setup for powder sample

Paper 55-739, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted March 9, 1955; made available for printing September 16, 1955.

H. R. LAIRD and M. J. ZUNICK are with the General Electric Company, Milwaukee, Wis.

Fig. 2. Present-day commercial diffraction equipment

the rectangular tube. Because of size limitations in the equipment, it is difficult to use a cylindrical side window counter tube with a gas path of much more than 1 inch. The advantages of the greater gas path, when dealing with shorter wave length radiation, are obvious. They are discussed more fully later on.

ELECTRONIC EQUIPMENT

In regard to pulse sizes produced, the Geiger pulses are of the order of a volt, while the proportional pulses are in the microvolt range. Hence a high gain preamplifier was built with a gain of about 25,000, shown in Fig. 7. Beyond the preamplifier, conventional ratemeters and scalers can be used, except that the initial circuits should have resolving times of the order of 1 microsecond.

When X-ray intensity is measured with a ratemeter and scaler, it is found that the X-ray intensity is directly related to the number of pulses per second. These pulses occur at random and hence are not evenly spaced in time. The ratemeter averages the pulse rate over a time interval which can be adjusted as desired. The scaler measures the time required to accumulate a specific number of pulses, which number can be set as desired. For a certain desired statistical accuracy, the necessary number of pulses or counts can be calculated. For example, in order to have a probable error of 1/2 per cent (%), about 16,000 counts must be registered.

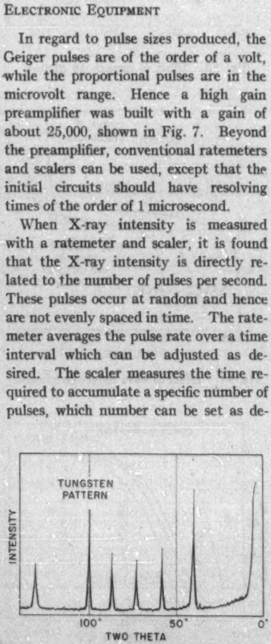


Fig. 3. Typical diffraction pattern

INTENSITY RESPONSE

In Fig. 8, the intensity response of a proportional counter is compared to that of a Geiger counter. Here the log of counts per second is plotted against the number of nickel filters placed to intercept the X-ray beam. Since the absorption curve is exponential, this relation should be linear. In this chart, the same tube, a rectangular one filled with argon and organic quench, was operated first as a Geiger counter and then as a proportional counter. First, it will be seen that the absolute number of counts per second is the same for both types of operation in the region where the two overlap. In regard to linearity, for Geiger operation, the curve appears linear up to about 1,500 counts per second (CPS) and the losses are 30% at 10,000 CPS. For proportional operation, the curve appears linear up to about 60,000 CPS and the losses are only 15% at 100,000 CPS, or 1.5% at 10,000 CPS. This means an improvement of 20 times in the reduced losses at 10,000 CPS for proportional operation.

It should be mentioned that the amount of departure from linearity shown in Fig. 8 for proportional operation is caused more by the limitation in resolving time of the electronic circuits than by tube resolving time. This over-all effective resolving time was found the same for both rectangular and cylindrical side window tubes, with any type of gas fill.

PLATEAU CHARACTERISTIC

In Fig. 9, curves 1 and 2 illustrate the differences in plateaus obtained for proportional and Geiger operation respectively. Both these are for the same tube, a rectangular one filled with argon and organic quench. The operating range is seen to be roughly 250 volts lower for proportional operation. In the same figure, curve 3 is for a rectangular tube filled with krypton with no quench, for proportional operation.

The question of the use of an organic quench in these tubes will be discussed fully in a later section. The effect of the quench on a typical plateau is shown in Fig. 10. The tube is a side window with xenon filling. Curve 1 is without and curve 2 is with quench. The use of the quench shifts the plateau upwards about 400 volts. Otherwise the plateaus are very similar.

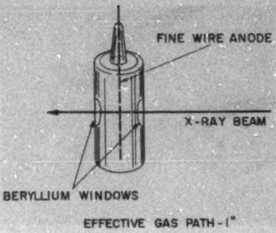


Fig. 4. Side window proportional counter

As mentioned previously, the size of the initial pulse before voltage amplification is inversely proportional to the wave length of the incident radiation. Hence the initial pulse produced by silver radiation, 0.563 Angstrom unit (A), will be about four times as large as the initial pulse produced by chromium radiation (2.29A.).. It is obvious that the voltage at the "knee" of the silver plateau will therefore be markedly lower than the voltage at the knee of the chromium plateau. Accordingly, a considerable shift of plateau with wave length is to be expected. A Geiger counter exhibits no such phenomenon because it pulse size, as explained before, is independent of the wave length of the incident radiation.

This shift of plateau with wave length for a proportional counter is illustrated in Fig. 11. The curves are for a side window xenon filled tube without quench. It will be noticed that the curves for silver and chromium have a greater slope than the curve for copper. This means that the pulse sizes for the copper radiation are more uniform than for either the

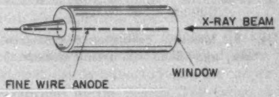


Fig. 5. End window counter tube

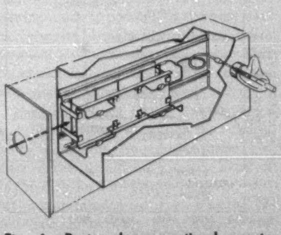


Fig. 6. Rectangular proportional counter, effective gas path 4 inches

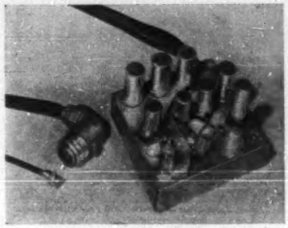
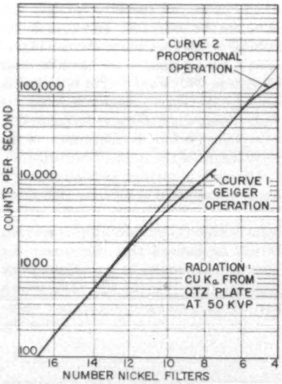


Fig. 7. High gain preamplifier unit

silver or the chromium radiation. Since the plateaus overlap for the two extremes, silver and chromium, it is possible to pick one voltage that will fall on both plateaus. Hence the tube can be used to cover a range of wave lengths without changing the applied voltage.



Fjg. 8. Intensity calibrations, rectangular tube, argon filled with organic quench

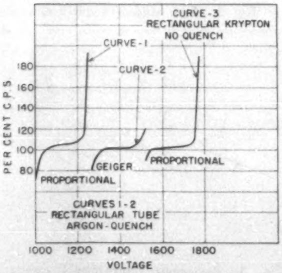


Fig. 9. Plateaus, copper radiation

EFFICIENCY VERSUS WAVE LENGTH

The question of the efficiency of counting over the range of wave lengths used in diffraction is a very important one. By efficiency is meant the percentage of the X-ray beam actually converted to counts. In Fig. 11 are shown the efficiencies of three different tubes. These efficiencies were calculated from the absorption coefficients of the gases, the path lengths and the pressures of the gases being taken into consideration. At the wavelength of chromium radiation (2.29 A), all the tubes were better than 99.5% efficient, so the curves all join at this point.

Curve 1 is for a rectangular tube filled with argon to 65-centimeter (cm) pressure. This tube is roughly 90% or more efficient from chromium to copper radiation (2.29 A to 1.54 A), but falls off to about 10% for silver radiation (0.563 A). Curve 2 is for a rectangular tube filled with krypton to 65-cm pressure. This tube is slightly more efficient then the argon tube down to 0.860 A, at which point the efficiency rises sharply. This sharp rise is due to the absorption edge of krypton gas occurring at this wave length. It will be noted that for silver radiation (0.563 A) the efficiency is about eight times that for the argon tube. Curve 3 is for a side window tube filled with xenon to 40 cm pressure. Although the pressure is less, and the gas path is only one-fourth that for the two rectangular tubes, the greater absorption of the xenon gas gives an efficiency almost equal to the tube filled with krypton, down to the krypton absorption edge. It can be seen that the choice of gas fill depends on the range of wave lengths which must be detected.

INTERFERENCE OF ANODE WIRES WITH NARROW BEAMS

The question has been raised concerning the interference of the anode wires with a very narrow beam such as is encountered in single crystal work. Obviously, since the wire necessarily crosses the center of the window, a marked decrease in intensity will be registered while scanning across a very narrow beam. This problem can be easily resolved by shifting the center of the tube slightly off the center line of the scanning goniometer. When this is done, the beam can be made to miss the wire entirely. By the time the beam approaches the wire, it will be no longer diffracted at all by the crystal.

TUBE LIFE AND USE OF QUENCH

To refer to a discussion by Curran and Craggs,[4] the life of a Geiger tube is of the order of 10^9 counts, while the same tube is estimated to have a life of 10^{12} counts when operated as a proportional counter. These figures refer to a counter tube containing an organic quench.

For the work reported in this paper cylindrical side window tubes were filled with argon, krypton, and xenon, both with and without an organic quench. Very similar results were obtained with the exception of the plateau shift mentioned previously. On the other hand, it was not possible to fill the rectangular tube with krypton, unless the quench was left out. With the quench, the tubes were always unstable and erratic.

In view of these facts all our proportional counters are now being built without quench. These counters have been

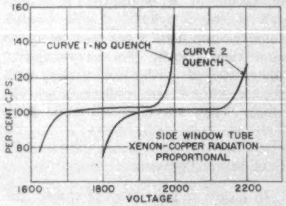


Fig. 10. Effect of quench on plateau

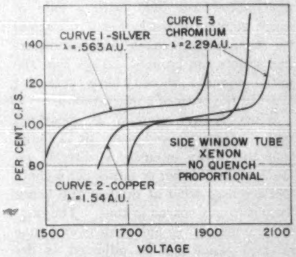


Fig. 11. Effect of wave length on plateau

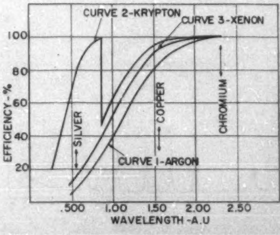


Fig. 12. Efficiency versus wave length

operated in continuous discharge for as long as an hour without changing any of the tube characteristics in the slightest degree. Hence under normal operating conditions, the life of these counters should far exceed the figure of 10^{12} counts for a tube with organic quench given in the foregoing. In fact, there is no predictable cause for failure in a tube of this type. Unfortunately, it is impractical to run actual life tests on such tubes. At a rate of 116,000 CPS, the counts per day amount to 10^{10}. For the low figure of 10^{12} counts, this would require one hundred days of test.

One other advantage to be gained by omitting the organic quench is the reduction of the temperature shift of the plateau. This point has not been investigated in great detail, but data on a rectangular krypton tube without quench indicate a shift of less than 25 volts in the plateau for a temperature rise from 20 to 40 degrees centigrade.

Conclusions

The proportional counter has been found to possess the following advantages compared to the Geiger counter with organic quench:

1. A life of at least 1,000 times that of the Geiger counter.
2. Reduction of the coincidence losses at 10,000 CPS by a factor of 20.
3. Elimination of the need for a quench with the following results: The elimination of the temperature shift of plateau occurring with the Geiger counter; the elimination of the damage due to overvoltage occurring with the Geiger counter.

References

1. The Proportional Counter in X-Ray Diffraction Work, A. R. Lang. *Nature*, London, England, vol. 168, Nov. 24, 1951, p. 907.

2. Wavelength Resolution of X-Ray Proportional Counters, A. R. Lang. *Proceedings*, Physical Society, London, England, A, vol. LXV, 1952, p. 372.

3. The Proportional Counter as a Low Background X-Ray Detector, D. J. Knowles. *Abstracts of Tenth Annual Pittsburgh Conference on X-Ray and Electron Diffraction*, Mellon Institute, Pittsburgh, Pa., Nov. 6–7, 1952.

4. Counting Tubes, Theory and Applications, S. C. Curran, J. D. Craggs. Academic Press, New York, N. Y., 1949, p. 41.

The Seattle-Victoria Radio System

R. E. KISTLER
FELLOW AIEE

ol. 74
. 1
955

THIS paper presents a description of a radio system now operating between Seattle, Wash., and Victoria, British Columbia. The radio system employs the Lenkurt *72A* equipment, the predecessor of the Lenkurt *72B* radio described in an unpublished paper by Harp and Kebby of the Lenkurt Company. This application of *72A* radio makes use of *33C* Lenkurt carrier with a capacity of 24 channels in comparison to the Lenkurt-*45*-class carrier used with the *72B* radio and having a capacity of 72 channels. The base band of the *72A* is 160 kc as contrasted to the 300-kc base band of the *72B*.

The application of this type of radio system with its associated carrier provides a typical example of a situation where point-to-point radio has been used to supplant an older type of facility. The Seattle-Victoria circuits were carried over a rather circuitous route of 152 miles consisting of an underground cable from Seattle to Everett, an open-wire carrier lead from Everett to Bellingham and an open-wire lead with several intermediate submarine cables between Bellingham and Victoria; see Fig. 1.

This route included an 8-mile international submarine cable owned jointly by The Pacific Telephone and Telegraph Company and the British Columbia Telephone Company. This cable has required rather frequent replacement due to armor wire corrosion in the swift tidal current. The original cable was replaced in 1928 and again in 1940. When it failed and was repaired in 1950, the condition of the armor was such that it was concluded that the cable could not be repaired again. It was estimated that the cable would last only from 2 to 4 years. Actually, by the time the radio was put in operation, there were only two good pairs remaining out of six in the cable. The cable has since been abandoned and as much salvaged as possible.

Planning

Different plans were studied for the replacement and reinforcement of this route including cable plant as well as radio. Since the circuits from Victoria all terminate at Seattle there was no necessity of retaining the route via Bellingham. Cost studies proved the radio plan to be more economical and flexible. A 2-link route from Seattle to Victoria with a repeater at Mt. Blyn was found to be practicable. However, a 3-link route with an additional repeater at Port Angeles was adopted having a total length of 95 miles. By adding terminal equipment at Port Angeles, relief and diversification were also afforded between Seattle and Port Angeles.

Traffic requirements over this route showed the need for some type of light route system as contrasted to a *TD-2* radio system with its capacity of 600 circuits per channel. Present requirements were 16 message circuits, and forecasts indicated that this would increase to 24 circuits in about 10 years. There were no immediate requirements for broadcast or television circuits. Therefore, various systems with capacities of 24 to 36 message circuits were investigated.

Pulse-type systems were compared with frequency division multiplexing systems for this route. The frequency division multiplexing systems were thought to be more advantageous for this application because of their flexibility and conformity with wire and cable carrier systems. At Seattle it was necessary to locate the antenna and radio equipment at East Office, while it was desirable to locate the carrier equipment at the Seattle Toll Office, a distance of about $1^1/_2$ miles. The use of radio employing frequency division multiplexing made it possible to transmit carrier circuits straight through the radio terminal without demodulation. This could not be done with pulse transmission using standard wire plant.

Experience both in the United States and Canada has indicated the desirability of utilizing frequencies in the low part of the ultrahigh-frequency band or in the very-high-frequency band for over-water transmission paths, such as the 22-mile section between Port Angeles and Victoria, in order to minimize fading difficulties. After considering various possi-

---

Paper 55-752, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted January 6, 1955; made available for printing August 12, 1955.

R. E. Kistler is with The Pacific Telephone and Telegraph Company, Seattle, Wash.

The author wishes to acknowledge the co-operation of B. R. Tupper of the Northwest Telephone Company who contributed information on the installation at Victoria.

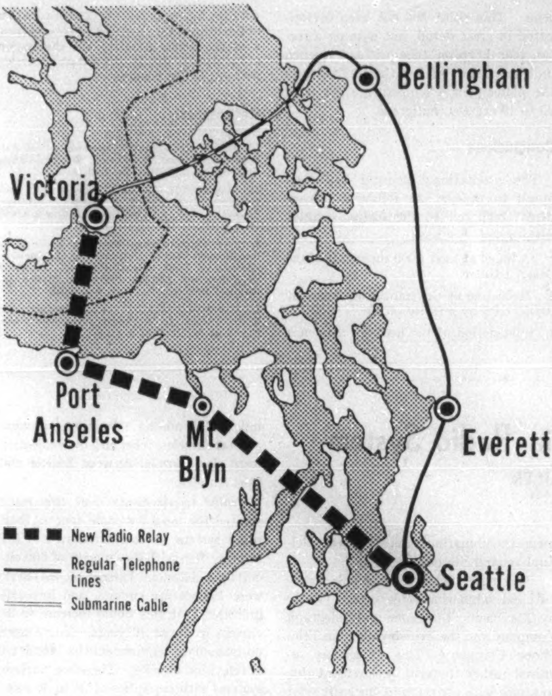


Fig. 1. Seattle-Victoria routes

bilities, it was decided to utilize frequencies in the range 450 to 460 megacycles (mc) for the Port Angeles–Victoria section. The Northwest Telephone Company was licensed by their Department of Transport to transmit at 454.25 mc and the Pacific Company was licensed by the Federal Communications Commission to transmit at 459.25 mc. The 450-mc transmitter inputs are limited to 40 watts with normal outputs of 8 to 12 watts.

However, the frequency band available in this range under Federal Communications Commission allocation for point-to-point service was not wide enough to accommodate the assignments required for the remainder of the route. Accordingly, frequencies in the range 890 to 940 mc were used for the Seattle–Mt. Blyn and Mt. Blyn–Port Angeles sections. The Pacific Company was licensed by the Federal Communications Commission to use frequencies of 890.5, 892.5, 921.5 and 932.5 mc. The link between Seattle and Mt. Blyn is 41 miles in length, partly over water and partly over land. The Mt. Blyn–Port Angeles link is 22 miles long and entirely over land. The 900-mc transmitter inputs are limited to 20 watts with normal outputs of 4 to 6 watts.

Antennas

Map studies were made of various routes for the system. When the intermediate repeater point was chosen between Seattle and Port Angeles a visual check was made of the line-of-sight path at night with the use of lights. The antenna was located to provide suitable clearance over the path as indicated by the light tests. However, the facilities for making light tests were not adequate to determine accurately the clearance over intervening obstructions. Consequently, it was found that less than 0.6 Fresnel zone clearance at 4/3 earth curvature existed in the path between the final tower locations in the Mt. Blyn–Port Angeles section, because of some intervening trees. The resulting additional path loss from this obstruction was approximately 10 decibels (db) and, while not serious under normal conditions, it subtracted this amount from the fading margin. No 900-mc path loss testing equipment was available at that time to make actual path loss tests.

A clear radio path does not exist between the central office building at Port Angeles and the central office at Victoria because of building obstructions at the Victoria end. The Victoria radio terminal was therefore installed in a penthouse on the Bank of Toronto Building, with antennas attached to the side of the penthouse, to give a clear path to Port Angeles; see Fig. 2. Approximately first Fresnel zone clearance exists over the Strait of Juan De Fuca, assuming 4/3 earth curvature. The carrier terminals were installed in the Victoria telephone building a block away and connected to the radio equipment by cable pairs.

A 70-foot self-supporting steel tower was erected on the roof of the Port Angeles building to carry antennas toward Victoria as well as Mt. Blyn; see Fig. 3. At Mt. Blyn a 70-foot triangular wooden pole structure was erected to support antennas in both directions; see Fig. 4. At Seattle East Office an existing steel platform terminating the Seattle-Portland *TD-2* microwave radio relay system furnished a suitable location for the antenna toward Mt. Blyn; see Fig. 5.

The antenna at Victoria consists of two 6-foot octagonal-shaped reflectors provided by the Radio Corporation of America. At Port Angeles, two Andrew Company *117*-type double 6-element Yagi antennas are used toward Victoria and located approximately 50 feet up the tower. A single 10-foot parabolic antenna is located at the top of the Port Angeles tower toward Mt. Blyn. At Mt. Blyn single 10-foot parabolic antennas are used in each direction. At Seattle a single 10-foot parabolic antenna is installed toward Mt. Blyn. The parabolic antenna mentioned were made by the Andrew Company. Gas-filled 7/8-inch coaxial lines of 50 ohms' impedance are run to the antennas in all cases.

Buildings

The only site to be acquired was the Mt. Blyn repeater station. This was located on state forest land at an elevation of 2,000 feet. An easement for a tract 300 by 300 feet and the right of access over a forest service road were secured from the State Forest Board. Easements were secured for use of the road from private

Fig. 2. Victoria antenna

property owners where the road crossed their lands. The telephone company will share in the upkeep of about 4 miles of road. A 40-foot right of way for the construction of a joint power and telephone line which did not follow the road was also secured from the State Forest Board and other forest owners, and compensation was made for the trees cut on this right of way.

The repeater building at Mt. Blyn, the only building to be constructed, is of fireproof design; see Fig. 4. It is a 20-foot by 30-foot steel frame building with corrugated galvanized steel sides and roof erected on a concrete slab foundation with linoleum flooring. The interior finish is plasterboard and the entire building is insulated with 4 inches of rock wool. The building is divided into three rooms for equipment, heating furnace, and power plant.

ol. 74
. 1
955

Equipment

At Mt. Blyn the radio equipment consists of regular and stand-by *72A* terminals back to back, mounted on six low floor-supported bays; see Fig. 6. At Seattle East Office the *72A* radio equipment is mounted on two high bays; see Fig. 7. At Seattle Toll Office the *33C* carrier equipment occupies six high bays arranged on three double-sided bays; see Fig. 8.

The *72A* radio equipment brings the signal down to base band frequencies and back up to radio frequency again at each repeater point. This feature is used at Port Angeles to drop voice frequency circuits by means of carrier group-splitting filters.

Stand-by radio equipment is provided for protection against system failure and to permit routine maintenance without shutdown. Automatic transfer is provided by means of sensing units in the antenna circuit to determine proper transmitter output and by pilot tones transmitted to associated receivers to detect receiver failures. Signals indicating improper operation of the equipment are routed by means of a Lenkurt *51A* alarm and control system to the Seattle toll test board, where maintenance men may be dispatched to investigate trouble at Seattle East Office or the maintenance people at Port Angeles may be sent to Mt. Blyn, as the case may be. Victoria and Port Angeles alarms appear locally. In addition, a transmitted 6-kc signal is passed to Victoria and then retransmitted to Seattle to act as a loop alarm to notify both terminals of any complete loop failure. A 6-second delay is incorporated to take care of normal equipment switching or short-time abnormalities. Control of the Mt. Blyn repeater is also by means of the *51A* alarm-and-control system. Three types of orders may be sent to the Mt. Blyn repeater as follows:

1. Equipment transfer (routine).
2. Start engine generator (routine).
3. Order wire signal.

Alarms sent by Mt. Blyn repeater are as follows:

1. Open door.
2. Fire (high temperature).
3. Various equipment transfer alarms.
4. Power fail (engine start).

This alarm-and-order-control equipment operates in the 4-kc to 8-kc range below the message channels by means of transmitted tones which operate an associated tone receiver. An order wire is provided under 4 kc with 1,000-20 cycle signalling and with loud speakers at the terminals.

The *33C* Lenkurt carrier system is a 24-channel 4-wire system utilizing the frequencies between 10 and 135 kc. Three groups of 8-voice channels are stacked by means of group modulators to fill this frequency range. The basic group of

Fig. 3 (left). Port Angeles tower

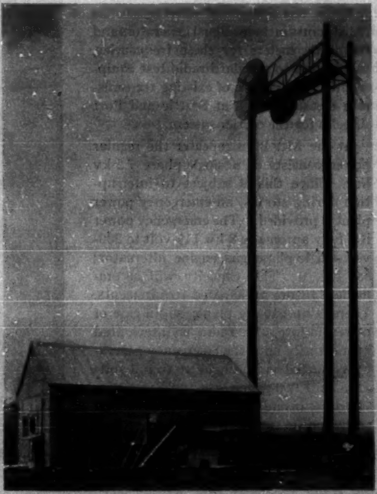


Fig. 4 (right). Mt. Blyn building and poles

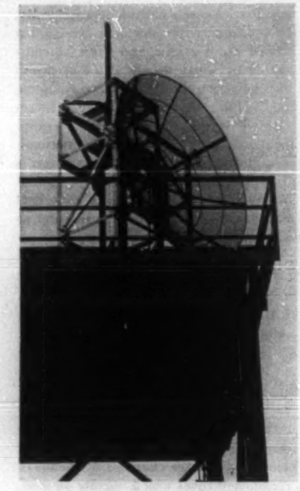


Fig. 5. Seattle East Office antenna

eight channels lies between 10 and 50, the second from 55 to 95 and the third from 95 to 135 kc. Single side-band suppressed carrier is used with out-of-band tone signalling. The radio inputs and outputs feed combining transformers to the three basic groups of carrier equipment which are located at Seattle and Victoria.

Radio test equipment provided at the three points includes signal generators and frequency meters for these frequencies, in addition to standard radio test equipment. Use is made of existing transmission measuring sets at Seattle and Port Angeles for the carrier system.

At the Mt. Blyn repeater the regular power consists of a single-phase 7.2-kv line. Since this is subject to interruption during storms, an emergency power plant is provided. The emergency power is a fully automatic 8-kw 115-volt to 230-volt single-phase gas engine alternator; see Fig. 9. This capacity will accommodate future anticipated requirements. There is no battery plant; so, in case of power failure, the radio is interrupted momentarily until the emergency generator is started and brought up to full voltage. A 1,000-gallon gas tank was provided to insure operation of the emergency power for a total of about 600 hours in case of extended commercial power outage occurring during winter storms.

Operation

There are 20 channels provided between Seattle and Port Angeles, eight of which are Victoria channels repeatered at Port Angeles, the other 12 terminating at Port Angeles. Between Port Angeles and Victoria there are 16 channels provided, eight of which, as mentioned, are Seattle channels repeatered at Port Angeles and eight of which terminate at Port Angeles and can be extended to Seattle over open-wire facilities. This provides a diversity of facilities for Port Angeles and Victoria and is of value when failures occur in either type of facility between Seattle and Port Angeles.

The system was placed in operation during June, 1953. Some troubles were experienced, as is usually the case in an initial installation. Some of the problems encountered and remedial measures adopted were as follows.

At the Mt. Blyn repeater the voltage regulation was poor and it was necessary to make some changes in the power system to maintain the voltage within proper limits. It was also necessary to maintain proper adjustment of the primary power transfer relay to avoid unnecessary starting of the gas engine when low voltages occurred. The final solution required the installation of a voltage regulator at Mt. Blyn.

Initially, the signalling circuits were of the "tone-on" type when busy, and any momentary interruption of the tone would ring on the circuit and interrupt calls. This was remedied by a relay modification changing the signalling to a "tone-off" basis, so that disturbances do not operate the signals and interrupt service.

The initial installation made use of a Lenkurt radio-frequency monitoring probe located in the antenna. The high loss of the probe when it became iced over gave an erroneous low-power output signal. This caused a transfer to the spare transmitter, which in time repeated the operation. This in turn caused hunting and gave a failure indication, and it was necessary to block out the monitoring circuit. To remedy the trouble, the radio-frequency monitoring probes were replaced with radio frequency sensing unit in the coaxial line (modified M.C. Jones station guardian) for low-transmitter alarm.

Certain relay-switching conditions were found that could leave a properly operating transmitter connected to a faulty receiver. This was corrected by a wiring change.

As mentioned previously, the Mt. Blyn–Port Angeles radio path was found to be 10 db high in loss because of low-clearance conditions. While this was not serious under normal conditions, it should be realized that the 10 db is lost from the fade margin. To recover this loss the 6-foot parabolas at Mt. Blyn and Port Angeles were replaced with 10-foot parabolas.

Noise-recording equipment consisting

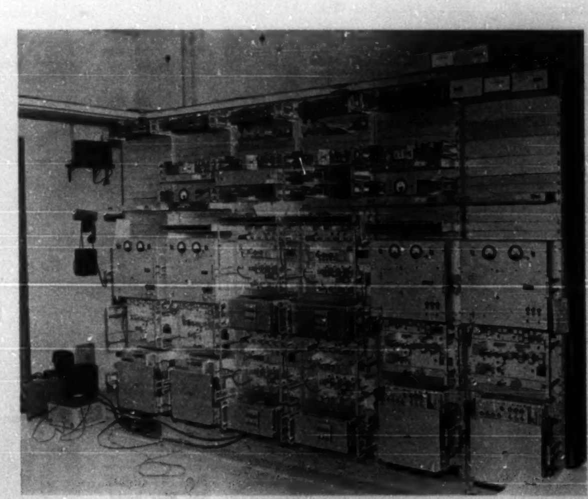
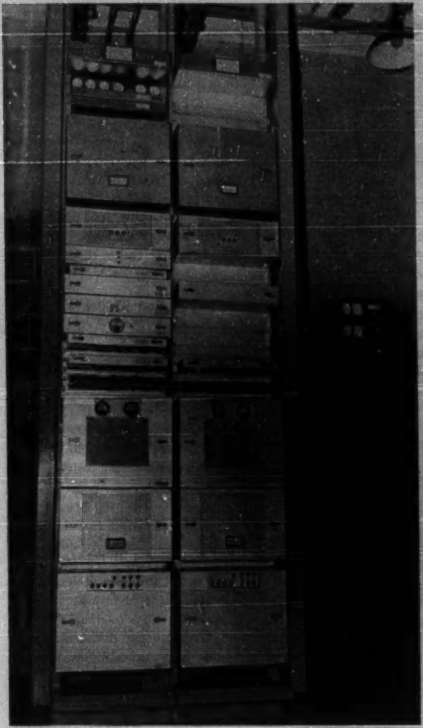


Fig. 6. Mt. Blyn radio repeater equipment

oL 74
. 1
955



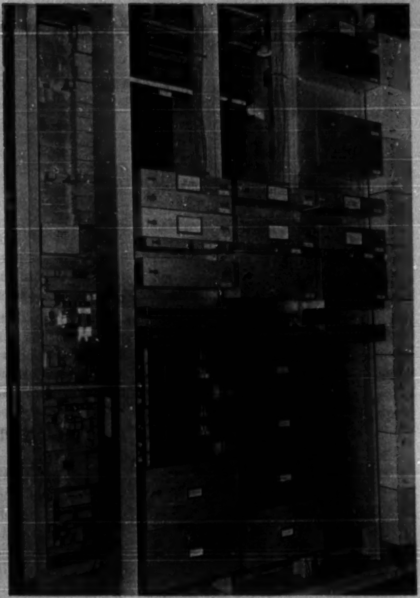


Fig. 7 (left). Seattle radio terminal equipment

Fig. 8 (above). Seattle carrier terminal equipment

of a General Electric Company recorder connected to the output of a *2B* noise-measuring set was installed for an extended test period on one of the message channels. This equipment recorded substantial noise increase during deep fades. However, during the first year of operation, there were no outages due to fading.

Interference on the 454.25-mc receivers at Port Angeles which caused severe noise on the service channel was identified as caused by aircraft radio altimeters (absolute terrain indicators). This was reduced to a satisfactory level by changing the polarity of the receiving antenna at Port Angeles from horizontal to vertical.

The snow conditions at Mt. Blyn make it impossible to make maintenance visits with ordinary automotive equipment. During the past winter, which was not exceptionally severe, the maintenance men were able to reach the top at all times with a 4-wheel-drive jeep equipped with chains.

During this first year of operation the radio facilities have been used on occasion to reroute circuits from other routes which were temporarily out of service, thus indicating the desirability of providing service by diversified types of facilities. On one occasion, the British Columbia Telephone Company lost their submarine cable between Vancouver and Victoria and were able to reroute a sufficient number of circuits via Seattle and Port Angeles to meet the emergency. On another occasion, the Pacific Telephone Company lost their wire line between Seattle and Port Angeles. In addition to the 12 direct Seattle-Port Angeles radio circuits, five circuits from Port Angeles to Seattle were routed from Seattle to Port Angeles over the radio system via Victoria. A Western Electric Company *C* carrier system was operated in the unused allocation for future *33C* channels 21 to 24 between Seattle and Port Angeles to provide three extra circuits.

Conclusion

The Seattle-Victoria radio system has performed satisfactorily during the first year of operation. Experience indicates that the *72A* Lenkurt radio equipped with *33C* Lenkurt carrier is well suited for systems of 24 circuits and the *72B* radio equipped with *45*-type carrier should be very suitable where 72 circuits are needed. The use of this type of radio system to supplement an existing land line route provides diversification and protection against complete failure, which is a highly desirable feature where adequate reroute facilities are not available. Dual equipment and automatic transfer have shown to date reliability exceeding 99.9 per cent.

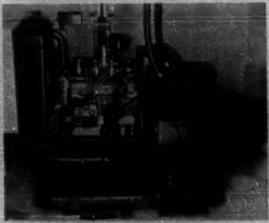


Fig. 9. Mt. Blyn emergency power plant

Development of a Wire Contact Relay

R. E. MARKLE
ASSOCIATE MEMBER AIEE

IN AN age where electronic devices are increasing in quantity and complexity, the electromechanical relay still plays an important part in various switching operations. Commercially available relays were formerly used in accounting machines but, with steady increases in machine speed and capacity, it became necessary to design an improved relay. The development of a wire contact relay for business machines applications is discussed herein.

Requirements

Two basic types of relays are required for business machines. In certain applications relays must make and break inductive loads of several hundred milliamperes but, in the majority of the applications, relays are required to set up the network matrix and carry current only. It has not been feasible to have one relay perform both functions because the demand for the first type of relay represents only a small percentage of the total number of relays required in an accounting machine.

The most important relay requirements are fast switching times, high reliability, and long life. These and other specific requirements are listed in the following:

1. Fast operate and release time.
2. Long life.
3. Must be compact and operate on low current.
4. Must be pluggable.
5. Nominal cost in mass production.
6. A variety of contact combinations in a modular structure.

The card capacity and life of the accounting machines established the required relay switching speeds as well as the desired number of operations. An operate time of 4 to 6 milliseconds and a release time of 4 milliseconds are required for this application, while a life of 100- to 200 million operations without adjustments is necessary. In addition, the relay must be compact and operate on low current because of the large number of relays used in a machine (600 to 1,000 per accounting machine).

A pluggable construction with a combination of 4, 6, or 12 contacts provides efficient manufacturing and ease of service of both the relay and the machine, with a corresponding over-all reduction in costs.

Development of the Relay

Initial studies indicated that these objectives would be most readily met by the wire-spring type of relay, as shown in Fig. 1. The round wires form a microcontact construction (less than 10 grams of force) which is inherently simple and inexpensive, and produces relatively high contact pressure with a small force. Furthermore, the small mass of the wire contacts results in very little bounce at high closing speeds. It was found that a wire of the eutectic silver-copper alloy would give both the desirable mechanical and electrical properties. Double wires are used to increase contact reliability. To fulfill the no-adjustment requirement, the relay is designed with sufficient wire flex to compensate for wear and to provide adequate contact force even after several hundred million operations. The wires can be easily removed and replaced.

The wires are actuated by an insulating plastic bracket attached to the armature. Wear of the bracket was at first a serious problem because of the constant flexing and sliding of the wires. A nonpigmented urea-formaldehyde resin was eventually found to give the necessary wear resistance.

The original production model, Fig. 2, was designed with a *C*-shaped moulded phenolic-formaldehyde frame to enclose and protect the contacts and to minimize the number of parts needed. The fixed contacts of the same eutectic silver-copper alloy are moulded into the frame with their ends protruding to make contact with the receptacle. A guide stud is provided to aid in installing and to prevent damage to either receptacle or contacts.

The wire contacts are held in position in the relay and forced against the normally closed contacts by the action of the common inserts, and the wire holder can be adjusted to alter the contact force. The necessary 5 to 8 grams of force per set of wires is obtained with an 0.016-inch-diameter wire of 100,000-pound-per-square-inch tensile strength. To permit fast release, an 0.004-inch armature-to-core air gap is provided and, in addition, a nonmagnetic separator of 0.003-inch-thick stainless steel is welded on the magnet yoke at the free end of the armature.

Present Production Model

After a few years of experience with the original production model, further refinements were incorporated in the design of the present production relay. These changes resulted in a 400-per-cent increase in relay life and in a 40-per-cent reduction in manufacturing costs.

The present production relay is illustrated in Figs. 3 and 4. To increase stability of relay operating characteristics and reduce the air-gap tolerances, the armature hinge was converted from a pin-type pivot to a knife-edge design. The use of the knife-edge pivot makes it possible to maintain close tolerances on the armature-to-core air gap during manufacturing, and the increase in the bearing surface reduces wear and the subsequent change in release time.

Another modification in the design involved the damping system for preventing armature rebound, which is the primary cause of contact bounce. Since armature rebound during operation caused no contact bounce, it was possible to replace the inertia damper, Fig. 2, with a soft vinyl stop under the wire bracket; see Fig. 3. This damping system reduced relay manufacturing costs and provided a more stable damper.

To eliminate the adhesion of the wire bracket to the vinyl damper caused by the plasticizer at temperatures above 55 degrees centigrade, the damper is made of a synthetic elastomer known as Buna *N*.

The final modification of major importance was in the construction of the wire bracket. This part is now moulded directly on the armature, replacing a comparatively elaborate, riveted assembly. To prevent splitting and to reduce warping, the bracket is moulded in sections, each section containing one mechanical anchor for the moulding and operating two pairs of contact wires.

Paper 55-763, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted March 31, 1955; made available for printing September 7, 1955.

R. E. MARKLE is with the International Business Machines Corporation, Endicott, N. Y.

Major credit for the development of the wire contact relay is due C. D. Lake and M. Fiehl of the International Business Machines Corporation.

Vol
Pt.
195

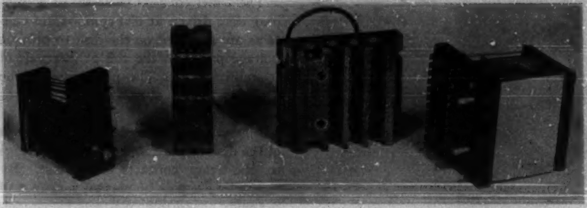


Fig. 1. Development models of wire contact relay and receptacle. Relay at left has 4 poles, relay at right has 12

Relay Receptacle

The relay is designed for easy replacement in the field by providing a receptacle, as shown in Fig. 5. This also permits complete cabling of the machine before installing the relay. A high impact-strength plastic was used initially, but was later changed to Hycar for greater resiliency.

A contact force of nearly 250 grams is provided between each terminal spring and its relay insert. This assures proper electric contact and helps to hold the relay in place. In addition, as a positive safeguard against loosening under vibration, the receptacle and the relay are secured together to the machine with a clip which engages recesses in the relay frame moulding.

Each terminal spring is made of beryllium copper, and is formed with two female sockets to facilitate wiring interconnection on the machine. To provide a solderless joint, a sleeve is crimped over the wire and thrust into the terminal spring as shown.

Fabrication and Subassembly

A special-purpose machine was built to form the wire contacts. Counterrotating die heads are pushed through the wires to straighten the wire, which is then cut by the machine to form the loop at the end.

The magnet circuit of the relay is fabricated by, first, projection-welding the extruded core to the formed yoke. After machining, to shape the armature fulcrum and establish the proper armature-to-core air gap, the assembly is chrome-plated; then the preformed stainless-steel residual plate is welded to the yoke. The magnet is completed by gluing the bobbin to the core.

The manufacturing techniques used for the relay frame combine high precision with relatively low cost. As seen in Fig. 6, each row of terminal inserts is stamped out as a slotted strip and moulded integral with the frame. After moulding, the strips are separated into individual terminals by sawing and milling. The contact points are chamfered, and any adhering plastic particles are removed from the inserts by tumbling and sandblasting. The frame is completed by assembling the damper, guide stud washer, and guide stud.

Final Assembly

The armature and coil springs are placed in the relay frame, and the yoke and wire holder are screwed in approximate position. The contact wires are added after setting the armature-to-core air gap.

Assembly is completed by adjusting the relay for proper operating characteristics. The relay is operated in a test receptacle, Fig. 7, with the armature vertical. Each terminal is connected to a test lamp indicating open or closed circuit. The tester is designed so that each contact may be checked separately, or all checked in series.

The yoke is first aligned so that all contacts "make" evenly. The wire holder, Fig. 3, is then adjusted while the armature is supported by a calibrated weighing device. The proper contact force is determined when the armature balances in mid-position, with no test lights on. When the wire holder is adjusted evenly, the force on each normally closed contact will be within the specification of 5 to 8 grams.

The force on the normally open contacts is controlled by the IN or OUT adjustment of the yoke relative to the free end of the armature. An 0.017-inch-thick feeler gauge is inserted between the yoke and armature and the yoke is moved until no contacts are made, as indicated by the test lights. As a check, an 0.010-inch-thick feeler is inserted between the yoke and armature, whereupon all normally open contacts should "make." These adjustments for contact force are easily made and have given uniform results.

The relay receives its final adjustment in a timing circuit, Fig. 7. At this time the damper is adjusted by means of a set screw, Fig. 3, to obtain the correct operate time and provide just enough clearance between contact wires and wire bracket to prevent armature rebound from disturbing the contacts.

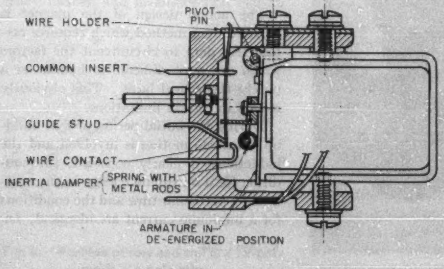


Fig. 2. Cross section of original production relay

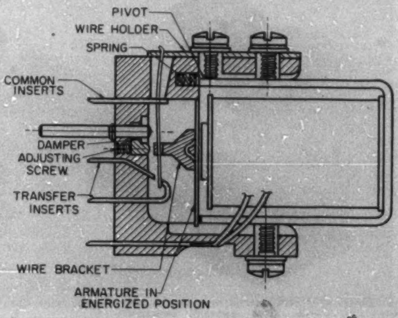


Fig. 3. Cross section of present production relay

ol, 74
1
55

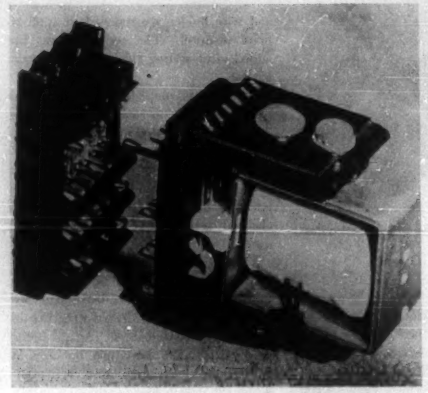
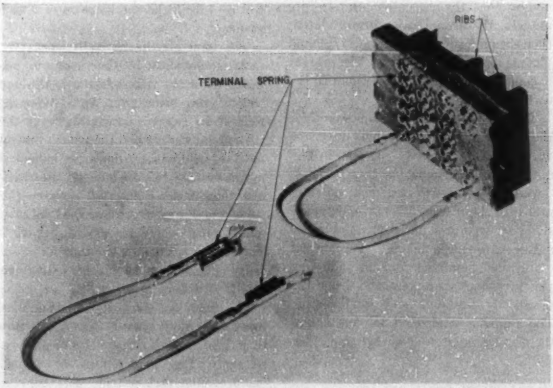


Fig. 4 (left). Present production receptacle and relay showing the 6-pole design

Fig. 5 (below). Relay receptacle and terminal springs



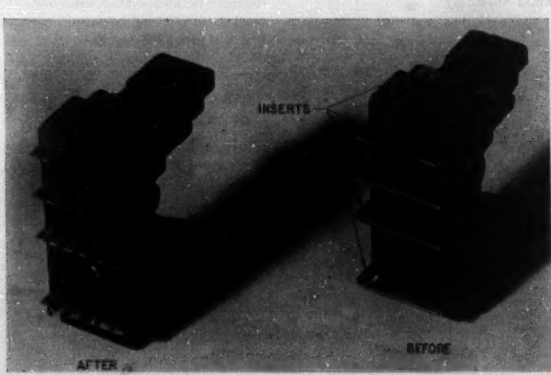


Fig. 6. Relay frame moulding showing processing of terminals

The timer consists basically of a spinning disk driven at 600 rpm. Cam contacts are mounted on the disk to provide voltage pulses to the coil at intervals of 50 milliseconds. The plate around the disk has 100 equal divisions corresponding to milliseconds.

A neon bulb mounted on the outer edge of the disk is connected through slip rings to the relay contacts under test. By connecting a pair of normally open and normally closed relay contacts in parallel, the bulb will be on at all times except during transit of the wire between contacts. The effect is a ring of light with a break at each side. If there is bouncing, additional breaks are seen. With the device, contact timing can be measured to less than 0.2 millisecond.

Magnet Details

To insure high machine reliability, the relay must be designed with the necessary safety factors. The relay must be able to operate on pulses of less than normal duration and on voltages considerably below normal. These safety factors are necessary because of the possible reduction in pulse width caused by tolerances in cam settings and improper sensing of the card, voltage regulation of the machine power supply caused by heavy pulse load, high ambient temperatures and coil heating, and circuit interaction.

The design attempt is usually to secure a 100-per-cent safety factor in time and current. With normal machine pulses of 10 to 12 milliseconds the relay must either operate in 5 or 6 milliseconds or on a 5- or 6-millisecond pulse, depending upon other circuit requirements. The 100-per-cent safety factor in current means that the steady-state current must be double the current at which the armature dynamically seals. This not only provides for the variance in supply voltage or other causes but secures the desired matching between the magnet design and the mechanical design.

The actual design of the magnet is handled by a method which requires certain test data to circumvent the factors which are most difficult to handle on a purely theoretical basis. This obviously results in certain limitations.

During the initial period of flux buildup, when no motion is involved and the inductance is nearly constant, the condition for a maximum energy in the magnet at a particular time and the conditions for a minimum current are identical; i.e.

$$R/Lt=1.253$$

where

Vol. 7
Pt. 1
1955

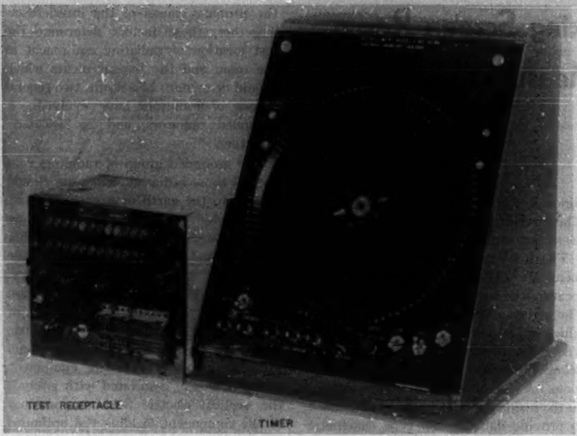


Fig. 7. Relay test receptacle and timer

L = the inductance of the magnet
R = the resistance of circuit
t = the time at which energy is to be a maximum

This simply states that, for maximum energy at a particular time, the relay must have a certain time constant. The problem is then handled experimentally by finding the optimum circuit conditions while keeping the operate time constant. In this manner it is possible to determine the optimum coil winding as well as the optimum magnet size.

Work that has been carried out in recent development combines the energy equations of the magnetic circuit, along with the mechanical equations, to permit the complete evaluation of a magnet design by means of an iterative process on a computer. The yoke and armature are made of cold-rolled *SAE 1010* steel, while the core is cold-rolled *SAE 1020*. Only the core is annealed. These steels are easily formed and machined and can be held to close tolerances. The magnet configuration reduces the effects of eddy current to less than 15 per cent of the total operate time.

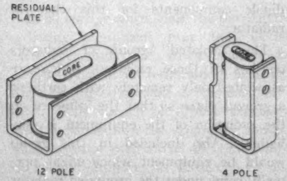


Fig. 8. Position of core and coil in a 12-pole and in a 4-pole relay

The high residual retentivity of the carbon steels is nullified, at the cost of increased relay current, by the nonmagnetic separator of stainless steel welded on the yoke and by the 0.004-inch air gap provided between the armature and core. An oval core is used to permit winding a greater number of turns without increasing the width of the relay. The bobbin is moulded with nylon. The coils are wound with an enamelled wire which is somewhat heavier than single enamel; they are encased in a 4-ply acetate yarn cover which is coalesced with acetone. Fig. 8 shows the yoke and coil assembly of two relay sizes.

Operating Characteristics

The relay is made with double throw contacts of 4, 6, and 12 poles. A variety of coils are used with as many as three coils on one relay. The operate time is plotted against power in Fig. 9. The release time is less than 4 milliseconds. Normal machine voltage is about 44 volts.

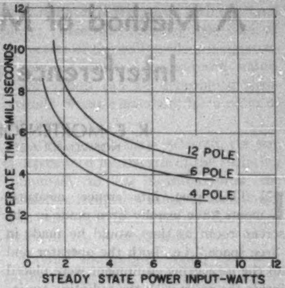


Fig. 9. Operating time characteristics of the wire contact relay with standard coils

The contacts will carry a 10-ampere steady-state current without serious heating, and may be used for short pulses as high as 40 amperes. Too high a current produces failure first in the receptacle by causing the springs to anneal and burn out. When contacts are used only for carrying current within these limits, relay life is a function of mechanical wear. After 100 to 200,000,000 operations, wear of the contacts, the wire brackets, the common inserts, and the wire holder cause a sufficient loss of contact force to result in unreliable operation. However, when contacts are used to make or interrupt current, relay life will be seriously affected because of the small mass of the wire contacts. Relays have been made for circuits in the microampere-millivolt range by a change in contact material and a modification in the contact design.

The inserts or fixed contacts are fabricated from a thrulay or special raislay construction of a noble alloy of gold, silver, and platinum known as number 1 metal. The wires are clad with the same material. This use of thrulay and cladding makes it possible to manufacture the relay with no additional costs except for the cost of the small amount of precious contact material that is used.

Conclusions

The wire contact design has provided relay that has fast operating characteristics with a relatively long life. In addition, this design has resulted in a compact component which can be manufactured at a very low cost. Several years of field experience have proved the relay to be satisfactory in accounting machine applications as well as many other special machines. Development work now in progress promises to produce a relay which has faster operating characteristics with less power, with longer life and improved contact performance, and all of this at a lower manufacturing cost.

Reference

A NEW GENERAL-PURPOSE RELAY FOR TELEPHONE SWITCHING SYSTEMS. A. C. Keller. *Electrical Engineering*, vol. 71, Nov. 1952, pp. 1007–12.

A Method of Making Screen Room Interference Measurements

K. E. MORTENSON
NONMEMBER AIEE

C. J. TRUAX
ASSOCIATE MEMBER AIEE

IN THE past, interference measurements have usually been made in the screen room as they would be made in "free space," i.e., both the operator and all the measuring equipment were placed inside the room with the radiating equipment. Under these conditions the resulting measurements had to be empirically correlated with free-space measurements. Furthermore, because of the presence of the operator and measuring equipment in the room, the results obtained were often difficult to reproduce. It was thus apparent that a more reliable method of making screen room measurements was desirable. Work to this end was initiated by the Coles Signal Laboratory by means of a study contract.

The objective of the resultant measuring method presented here was to eliminate all measuring equipment and operators from the room except for small wall probes and then essentially treat the room as a cavity (below resonance) with only the radiating equipment inside. Reliable interference measurements could then be made provided that the fields created inside the room could be properly detected and related to free-space field intensity measurements.

Philosophy of Measurement

An analysis of radiating equipment has indicated that the radiated fields from such equipment could be generated by a properly oriented system of electric and magnetic dipoles. Furthermore, if the size of the radiating equipment is very much less than the wave length under consideration, the system of dipoles can be reduced to one resultant electric dipole and one resultant magnetic dipole. These results are the basis of the screen room measuring procedure described.

The dipole moments of radiating equipment are determined by appropriately locating the equipment in a screen room. The screen room is treated as a cavity which is excited internally by dipoles whose frequency of excitation is kept below the lowest resonant mode of the room. A system of pickup devices located on the inside surfaces of the walls is used to provide data on the type (magnetic or electric), strength, and orientation of the dipole present. The pickup system consists of both loop and rod probes appropriately located and polarized so that components of the electric and magnetic dipoles present can be individually detected. In this manner, by proper rotation of the pickup system or the radiator, all the components of each type of dipole can be detected. The strength of the detected dipole components are measured and correlated with the fields that such components would produce in free space (possibly over a ground plane) at a specified distance. This correlation is accomplished by calibrating the probes, so that the free-space fields can be determined by applying a calibration factor to the probe outputs. All measurements for the grounded radiator were correlated to the vertical electric field as this was considered the optimum measure of interference for the particular applications of the Signal Corps. However, this system is not limited to the consideration of this field component only.

Method of Measurement

GENERAL

Consideration of Radiator Configurations

To obtain a measurement which will provide a meaningful indication of the interference effect produced by radiating equipment, two prime factors must be considered: 1. the arrangement or layout of the radiating equipment including all its associated cables and supports, and 2. the determination of which field intensity component or components are the primary causes of the interference. Together, these factors determine the test location of radiating equipment in the room and the measurements which should be made. As a result, two general classes of radiators have evolved: 1. grounded radiators, and 2. "isolated" radiators.

The grounded group of radiators consists of those radiators which are closely related to the earth or a ground plane. This includes equipments which are directly bonded to a ground plane either by an actual ground strap or by cable connections or equipment whose nearness to a ground plane influences the radiation from the equipment. An example of this class of radiator might be electric machinery mounted on a ground plane. For this group of radiating equipment which is closely associated with ground, the vertical electric field was selected as the component yielding the optimum measure of interference. This selection was made for several reasons: 1. No horizontal electric field component would be radiated along the ground plane; thus, only the vertical electric component exists; 2. Most communications equipments (particularly those of the Army) which might be interfered with employ a whip or vertical-wire type of antennas which obviously respond to the vertical electric field component, not the magnetic components; and 3. Existing field intensity measuring procedures and specifications generally treat the vertical electric field component.

Radiating equipments of the grounded class which are small compared to the wave length of measurement can be represented by a vertical electric dipole and a resultant horizontal magnetic dipole. The two horizontal electric dipole components (assuming three orthogonal components) are not considered, since their ground plane images produce opposing fields and, for the same reason, the vertical magnetic dipole component is not considered. Thus, the room pickup system must be capable of separately sensing one vertical electric dipole and two horizontal quadrature magnetic dipole components for this class of radiator.

The "isolated" group of radiators consists of those radiators which are associated only remotely with earth or a ground plane so that the influence on the radiation of the equipment is negligible. Also included in this group would be equipment which might normally come under the grounded radiator class but, because of the need to determine the interfering fields close to the radiating

Vol.
Pt. 1
195

Paper 55-737, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3-7, 1955. Manuscript submitted March 17, 1955; made available for printing September 16, 1955.

K. E. MORTENSON and C. J. TRUAX are with Rensselaer Polytechnic Institute, Troy, N. Y.

The authors wish to express their gratitude to the Suppression and Engineering Section of Coles Signal Laboratory for their initiation and sponsorship of this work; to all the members of the project who helped in its development; and in particular to Gerard Hasserdji, who was responsible for an appreciable portion of the ideas and the effort required to develop this method of measurement.

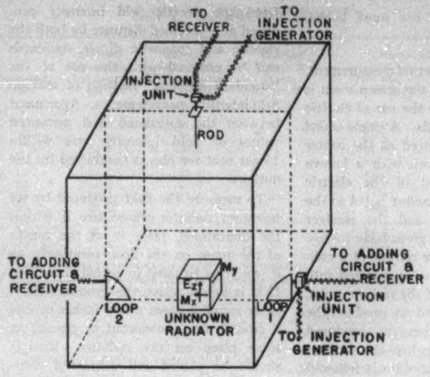


Fig. 1. Arrangement of pickup probes for making measurements on a grounded radiator

equipment relative to the nearest ground plane, other field components (e.g., the horizontal electric field component) must be measured. Examples of this class of radiator might be battery-operated equipment not associated with any ground plane, and bench type of electronic test equipment which might be grounded but which operates sufficiently close to interference-susceptible equipment so that all field components might be important. For this group of radiating equipment, any field component could be selected as the component yielding the optimum measure of interference, depending upon the particular type of radiating equipment to be measured and the typical type of equipment which must be protected from interference.

Radiating equipments of the isolated class which are small compared to the wave length of measurement will be represented by a resultant electric dipole composed of three orthogonal components and a resultant magnetic dipole also composed of three orthogonal components. Note that all possible dipole components must be considered in this case for the reason that there is no ground plane to cancel some of the components out.

In no matter which class the particular radiating equipment may fall, it is extremely important when making a measurement of its radiation to arrange the radiating equipment as it is normally used. This normal or typical arrangement includes both the grounding conditions and the position of all connecting cables and supports, as these can be an important portion of the equivalent antenna configuration of the radiating equipment.

The Room and its Pickup System

As the screen room serves principally to provide a noise-free volume of space, the shielding provided by it need only be effective enough to eliminate ambient noise in the room relative to the minimum noise fields to be measured. Thus, the room used for these measurements might be constructed of coarse screening or might require double-room or sheet metal construction depending upon the room location and the measurements to be made. As the room also serves as a conductive boundary for the radiation produced inside, the geometry of the room is important. A cube-shaped room or cavity was selected for the following reasons: 1. A room with flat surfaces is the most convenient for mounting equipment and pickup devices. 2. For calibration purposes, it was desirable to have the pickup devices and thus the walls equidistant from the vertical axis. 3. The length of the walls should be approximately equal to the height of the room (floor-to-ceiling spacing) for the optimum compromise between room sensitivity and room volume. Room sensitivity is defined as the ratio of the value of the field components (magnetic or electric) at the room boundaries to the value the same components would have with no boundaries present.

Because of the existence of a metallic boundary about the radiating object, the fields created at the pickup devices are different from those which would be created at the same pickup points if no room existed; e.g., the electric field (created by an electric dipole) in the room (5-foot cube) at the center of the ceiling is approximately three times the value created at the same point by the same radiator in free space with a ground plane. Likewise, the magnetic field at the center of the ceiling edges is approximately 14 times the no-room value. These values can be determined theoretically or experimentally for a room of any size.

With the room shape selected, the size is determined by the size of the radiating equipment to be measured and the measuring frequency range in accordance with the specified system limitations which are to be discussed.

Having the proper room, the next factor of primary importance is the selection and placement of suitable pickup devices on the inside surface of the room. Two types of pickup devices are used: an electric probe consisting of an insulated perpendicular rod and a magnetic probe consisting of a 90- or 180-degree shielded loop. The same electric probe is used over the entire frequency range but different loops must be used for different frequency ranges. Each pickup unit has its own injection circuit for the measurement of the induced voltage.

The location of some of the pickup units in the room are illustrated in Fig. 1 and the completed room is shown in Fig. 2. The setup illustrated is primarily for measuring interference from the grounded class of radiators but can be used for the measurement of some of the radiators of the isolated class. The locations and connections of the probes are determined by the need to sense each dipole component separately. In several instances there is more than one possible location for the pickup device. The particular pickup units used for each dipole component measured in both classes of radiators are itemized in the following.

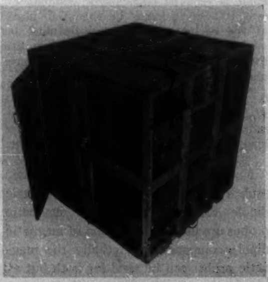


Fig. 2. External view of a typical shielded room showing output and injection cables

Grounded Radiators: Three Dipole Components at Center of Floor

1. Vertical electric dipole component E_z is detected by an electric probe placed at the center of the ceiling. This probe is sensitive to no other dipole components.

2. Horizontal magnetic dipole component M_y is detected by two floor or ceiling corner loops in the X-Z plane. Two loops must be used so that in adding their outputs (or paralleling them) the opposite-phased voltages induced in the two loops by the magnetic field of a vertical electric dipole will be cancelled out. These loops are not sensitive to the other magnetic dipole components.

3. Horizontal magnetic dipole component M_x is detected as M_y either by rotation of the radiating equipment 90 degrees into the M_y direction or by another pair of corner loops in the Y-Z plane.

Isolated Radiators: Six Dipole Components at Center of Room

1. Vertical electric component E_z is detected by an electric probe placed at the center of the ceiling as for the grounded radiators.

2. Horizontal electric dipole component E_x is detected by two opposite wall probes both in the X-Z plane placed at the center of the walls. The probe outputs must be subtracted. Two probes are used such that, if a vertical power cable is necessary (bench-type equipment), the probe-induced voltages it produces will be cancelled out.

3. Horizontal electric dipole component E_y is detected as E_x either by rotation of the radiating equipment 90 degrees into the E_x direction or by another pair of wall probes in the Y-Z plane.

4. Vertical magnetic dipole component M_z, is detected by two diagonally opposite corner loops in X-Y plane which passes through the center of the room. By adding the loop outputs together, the resultant output is produced solely by the M_z component.

5. Horizontal magnetic dipole component, M_y, is detected by two diagonally opposite corner loops (floor and ceiling) in the X-Z plane. The loop outputs are added. If a vertical power cable to the radiating equipment is required (bench-type equipment), two 180-degree loops in the X-Z plane at the center of opposite walls must be used in place of the diagonally opposite corner loops. This is done to eliminate undesired pickup from the power cable.

6. Horizontal magnetic dipole component M_x is detected as M_y by rotation of the radiating equipment 90 degrees or by similar loop pickup combinations in the Y-Z plane.

Note that in all cases the electric probes are used for detection of electric dipole components and the magnetic probes are used for detection of magnetic dipole components. Actually the magnetic probes can be used for both but at low frequencies they are much less sensitive than the electric probes for the detection of electric-dipole components.

PROCEDURE

Grounded Radiators

To accomplish the actual measurement, the pickup system of the screen room is initially calibrated by the use of electric and magnetic standards. A single-ended dipole standard is placed at the center of the floor and driven with a known voltage. The output of the electric pickup dipole on the ceiling is fed to the input of a receiver and the receiver indication noted for a given base voltage on the standard. The output of a calibrated signal generator is injected into the injection circuit of the pickup probe, and the drive required to produce the same output indication (as produced by the induced voltage) at the receiver is recorded. This procedure is followed through the frequency range desired. From the known drive voltage, capacitance and effective height of the standard and the frequency, the electric dipole moment can be determined. The resulting vertical electric field, at any distance, produced by this dipole moment can be computed from field equations. From the calculations, the dipole moment or the electric field intensity at a specified distance per injected voltage can be determined and plotted versus frequency for a calibration curve.

The probes used for measuring the magnetic dipole moment can be calibrated in a manner similar to that used for the electric pickup probes. As was mentioned previously, the set of loops on the floor or that on the ceiling may be used, depending on the size of the radiator and sensitivity required. It is therefore necessary to calibrate each set of loop probes.

A small standard shielded loop is placed at the center of the floor in the plane of the loop pickup probes to be calibrated. The output of the magnetic pickup loops is fed to the input of a receiver and the receiver indication noted for a given base voltage on the standard. A calibrated voltage is applied to the pickup loop injection circuit and the drive required to produce the same receiver indicated is recorded. The procedure is repeated for a number of frequencies in the range desired. Knowing the base voltage, area, and impedance of the standard at each frequency, the dipole moment and vertical electric field at a specified distance can be determined. From these data, a calibration curve for each set of loops can be plotted similar to that described in the calibration of the electric pickup probe. The calculated values of the free-space electric field intensity produced at a specified distance by both the electric and magnetic dipole standards can be checked with the use of the "standard antenna" method of making field intensity measurements. Agreement between the calculated and measured values of field intensity was within 15 per cent for checks conducted by the authors.

To measure the field produced by an unknown radiator whose size is within the limitations, place it at the center of the room on the floor identically as it would be situated in actual operation. If it is raised above the ground with a cable running down a few inches or has a ground strap connected to ground at some place on the radiator, then it should be placed and connected identically in the screen room. As has been mentioned, there are three dipole moment components of the grounded radiator that will produce a vertical electric field, namely, the vertical electric dipole moment and the two horizontal magnetic dipole moments. It is therefore necessary to measure the vertical electric field produced by each of these and then combine them.

To determine the vertical electric dipole moment and its corresponding field, measure the injected voltage in the electric pickup probe required to produce the same indication at the receiver as was caused by the field from the radiator. From the calibration curve, the calibration factor can be found and the vertical electric dipole moment and vertical electric field determined.

In a similar manner measure the injected voltage in the magnetic pickup loops required to produce the same indication at the receiver as the field from the radiator. With the use of the calibration curve, the horizontal magnetic dipole moment and its corresponding vertical electric field at the specified distance can be determined. Rotation of the radiator 90 degrees permits the determination of the electric field from the other horizontal magnetic dipole moment.

Without regard to phase, add algebraically the resultant of the two vertical electric fields produced by the horizontal magnetic dipole moments to the electric field due to the vertical electric dipole moment. This will give the maximum vertical electric field possible at the specified distance from the radiator. This maximum field is believed to be the best indicator of the radiation field because the phase between components and the resulting field pattern are unknown.

Table I. Comparison of Shielded Room and Field Measurements on Unknown Radiator

Frequency, Mc	Injected Voltage M_x; Top Loops, Db per Mc*	Injected Voltage M_y; Bottom Loops, Db per Mc	Injected Voltage E; Probe, Db per Mc	Corresponding Electric Intensity at 5 Feet			Total Intensity by Room Measurements, Db per Mc per M	Total Intensity by Field Measurements, Db per Mc per M
				From M_x, Db per Mc per M	From M_y, Db per Mc per M	From E, Db per Mc per M		
1.5	78.0	63.5	†	65.0	35.9		65.0	
3.0	75.0	60.0		61.8	32.4		61.8	
5.0	64.7	56.0		51.6	28.3		51.6	
5.0	61.5	53.5		54.4	30.7		54.4	
10	53.5	58.0		45.8	34.5		46.0	
15	53.0	61.0		46.9	39.2		47.5	
20	56.5	65.5	33.5	51.8	45.3	48.5	56.6	56.0
25	56.0			51.1			51.1	60.0
30	63.0	79.5	47.0	59.6	60.1	62.0	66.4	67.0
35	69.7	86.0	52.5	69.5	68.9	68.0	76.4	74.0
40	75.5	90.0	48.5	75.5	69.7	76.5	82.6	71.0

* db=level above 1 microvolt per megacycle per meter.

† Blank spaces indicate where readings were below receiver noise level.

Isolated Radiators

The isolated radiators basically consist of a self-contained unit but can be extended to include such items as bench equipment with a power cord, as was set forth previously. The self-contained radiators which have no cables or power cords, such as battery-operated equipment, are measured in a manner quite similar to that used for the grounded radiators. Because the standards and the unknown are placed at the center of the room it is necessary to use different sets of loops and probes.

For the self-contained radiators, those not receiving or delivering cabled power, the diagonally opposite loop probes and opposite wall electric probes are used. The diagonally opposite loops are used to eliminate the magnetic field pickup due to the horizontal electric dipole moments. Two opposite electric dipole probes with their outputs subtracted are used instead of one, so as to eliminate any cable effects in calibration.

The first step in the calibration is to place a balanced dipole standard in the center of the room in line with the two opposite electric pickup probes. With the use of a known drive voltage on the standard, obtain a calibration curve of dipole moment per injected voltage versus frequency by the same method as was described for the grounded radiators. To calibrate the magnetic pickup loops, a standard shielded loop is placed at the center of the room in the plane of the pickup loops. The calibration curve of magnetic dipole moment per injected voltage versus frequency is then determined as previously described.

Since a radiator of this type may be oriented in several different positions, and each dipole moment component may produce interference it is necessary to measure each of these components by the method described for the grounded radiator. The radiator is placed at the center of the room and rotated as required to determine each moment. The field desired at any frequency and distance may be determined from electromagnetic field equations.

For other types of isolated radiators previously mentioned, (such as bench-type equipment) it becomes necessary to make an analysis of the radiation possible and the field component(s) desired. From this, the pickup probes may be located to eliminate undesired components and to obtain an injected voltage proportional to those desired.

Example of Measurement

Interference measurements were taken on a battery-operated spark generator in the screened room and compared to those taken by the standard antenna method of measurement in the field.

In making room measurements the radiator was located at the center of the floor and bonded with the use of ground straps. The procedure followed has been outlined previously for measurements on grounded radiators. In the field, the radiator was bonded in a similar manner to that in the room. The electric fields obtained by the standard antenna method are compared in Table I to the calculated values obtained using the screened room and its predetermined calibration factors.

Measurements by the standard antenna method were not obtainable below 20 megacycles (mc) because of the low sensitivity of the system used. However, good agreement was obtained for readings where comparisons could be made, except at 25 mc. It was felt that the accuracy of measurement at 25 mc per second, both in the room and outside, was poor because of low sensitivity due to abnormally high receiver noise level. This would account for the poor agreement at this frequency. All other readings appear to be within the limit of accuracy. The shielded room measurements should be higher since the components are combined to give the maximum possible electric field intensity without regard to phase.

The good agreement of the measured electric field intensities gives proof of the usefulness of the shielded room for making interference measurements. Besides giving the total field intensity, it also indicates which components are the primary sources of interference.

Limitations

The methods of measurement outlined have certain limitations. First, the size of the radiating equipment must be small compared to the wave length of measurement so that the equipment possesses a simple dipole radiation pattern. It was determined that a radiator whose largest dimension was less than approximately one-tenth the wave length of measurement would give satisfactory results. Secondly, the frequency of measurement should be below the lowest resonant mode of the room. To employ the methods outlined here for making screen room measurements, it is felt that the frequency of measurement should be less than half that of the lowest resonant mode. Thirdly, the size of the radiator and pickup probes must be small in comparison to the room size. This limitation is required because the measurement is based upon the concept of point sources and probes. It was found experimentally and analytically that the ratio of radiator size to room dimension should be less than one-third. The radiator size limitation is determined by the degree of inaccuracy which can be tolerated in the measurement. Since this system of room measurement is precise only for point sources (infinitesimal radiators), any deviation from such sources introduces errors into the measurement. Thus, if the calibration factors which were based on radiators whose dipole moments were concentrated at the floor center (i.e., a point source-grounded type of radiator) were used when measuring a radiator whose size is an appreciable portion of the room dimensions, then considerable error will exist in the results obtained. This error results from the fact that unlike the standard dipoles used for calibration the dipole moments can be distributed uniformly throughout the radiator or they can be concentrated at a point different from the point where the room calibrating standard was placed. It was found best to assume that a uniform dipole distribution existed for large radiators, unless known to be otherwise, and to correct the calibration factors accordingly. Even with this correction in the worst case, when the electric dipole moment is concentrated at the point nearest the electric pickup probe, measurements on a grounded type of radiator one-third the room dimensions can be off by as much as 6 decibels (db). However, this extreme case is very rare. Similar magnetic dipole measurements can be made to within 3 db. It should be noted that for a given allowable tolerance of measurement for the grounded type of radiator the radiator can extend much more in the horizontal plane than in the vertical direction since the electric pickup probe used is mounted at the center of the ceiling and there are two oppositely placed loop probes. For further details on the accuracy of this system of measurement when large radiators are under consideration, the reader is referred to reference 1.

Conclusions

This screen room method of making interference measurements, which is approximately equivalent to making free-space field intensity measurements, should prove particularly useful when forced to make measurements in electromagnetically noisy regions. Furthermore, a measurement of the radiator equivalent dipole moments provide information from which the maximum resultant field intensity at any distance can be computed for any radiator arrangement. It is recommended that interference specifications be established with the use of this method of screen room measurement as a standard. Finally, it should be noted that it has been the purpose of this paper to set forth the general method of making comprehensive measurements of radiation in a screen room. Many specific items such as use at higher frequencies, placement of the radiator and probes, methods of electrically combining the individual fields to obtain a resultant, construction of the room, and typical ground conditions may require further study depending on the accuracy required, type of radiator to be measured and frequency range desired. It is hoped that the work presented here will initiate use of this method and stimulate further development.

The details of this method have been published in the quarterly reports of a Signal Corps contract.[1] Reports numbered 5 through 9 deal most directly with the subject presented in this paper. All pertinent references are included in these reports.

References

1. AN INVESTIGATION OF PROBLEMS ASSOCIATED WITH BROAD BAND INTERFERENCE MEASUREMENTS, U. S. Signal Corps contract DA 36-039SC-42488; file no. 26385-PH-52-91 (4502), Washington, D. C., 1954-55.

◆

Discussion

Harry R. Meahl (General Electric Company, Schenectady, N. Y.): It is evident that much intelligent effort has been put into the work reported in the paper, but in my opinion the author has not obtained enough quantitative data to provide a basis for using this method as a standard.

It appears to me to be useful over the frequency range of 0.1 to 50 megacycles per second for equipment 18 inches on an edge maximum.

Those concerned with testing small devices for electromagnetic radiation will do well to make a careful comparison of this method with their present method of test. It may prove superior.

K. E. Mortenson and **C. J. Truax**: A standard procedure for making screen room measurements has long been needed. Such a procedure or method should employ a fundamental rather than empirical approach and should provide some basic data instead of some arbitrary indication. It is felt that the method outlined in this paper fulfills these requirements, at least for the measurements on small radiators. Furthermore the use of the room itself provides a consistent ground plane when used with the grounded class of radiators, a condition which is not always true for field measurements and which must be true if reproducible results are to be obtained.

It has been suggested that further quantitative data be taken to provide a basis for acceptance of this method as a standard. Sufficient data have been taken to verify the fundamental soundness of this method, aud however, further data would be useful to substatiate the measuring tolerances which were derived theoretically and only partly verified experimentally. In addition, correlation factors between data taken in conjunction with a good ground plane (copper floor of the room) and verious ground conditions, as the particular equipment might be used, would definitely prove useful.

In using this method it is important to realize that the useable frequency range, room size, and the equipment size are all interrelated such that the statement made in the discussion regarding frequency range and size refers specifically to the 5-cubic-foot room. A 10-cubic-foot room, for example, could be used up to about 35 megacycles for equipment 3 feet on an edge.

Finally, if sufficient use of this method is made by those concerned with making interference measurements, then ample data on all points will be available from which an appropriate judgment of this measurement as a standard can be made.

Operation of a Nuclear Power Plant on an Integrated Electric System

N. E. WILSON
MEMBER AIEE

Synopsis: A nuclear power plant of the pressurized water-moderated and water-cooled type is able to meet the requirements of a central station facility for load changes and for picking up load. It may be used for system load control the same as for conventional plants. The negative temperature coefficient of reactivity characteristic adjusts reactor power level to meet power changes of the steam plant, and provides plant stability. The automatic temperature control system adjusts the control rods to compensate for reactivity changes in the reactor caused by fission products.

The plant may be shut down when not needed and returned to service at will. Approximately 1 hour is required to start up the plant and bring it to full power if it is hot at the time of start-up. About 3$^1/_2$ to 4 hours are required if the plant is cold. The major difference between this plant and a conventional one with respect to shutdown is in the heat produced by radioactive decay of fission products, which requires that there be provision for removal of this heat for some hours after shutdown.

A new core for a reactor plant can be designed for a specific number of hours of operation during which it is fully flexible with respect to load changes and start-up after shutdown. The core may be replaced at this time, or it may be used to provide base load until operation is no longer possible at design temperature because of fuel depletion.

ol. 74
. 1
955

WITH the successful demonstration of the use of heat of nuclear fission for power for ship propulsion, and its imminent use for commercial electric power production, many questions arise concerning the operation of this novel heat source. It is the intent of this paper to answer some of these questions. The discussion is directed specifically to the use of nuclear energy in central station service, and the questions posed are believed to be of interest to electrical utility engineers.

Requirements of Heat Source for Central Station Use

A nuclear reactor plant of the type discussed in this paper, i.e., one that is moderated and cooled with pressurized water, produces as its end product heat in the form of steam with which to power a conventional turbine. Since this steam is of a conventional variety, and not "hot" in the radioactive sense, the only aspect of the power generating plant to be concerned with is that portion dealing strictly with steam production. Of concern, then, is the part of the plant corresponding to the steam generator through to the main steam header of a conventional plant.

What are some of the operational characteristics of a steam generator which the operators of central station facilities have come to regard as either necessary or desirable? In broad terms, the requirements may be stated as follows:

1. It must be possible to start the steam generator when it is wanted, without excessive waiting periods.
2. It must be possible to load it to full capacity as fast as the electrical system requires, and to have it absorb transient load changes of the speed and magnitude required by the system.
3. It must be capable of surviving the emergency of having all of its output suddenly cut off, and yet of being restarted quickly if required.
4. It must be capable of being shut down when its output is no longer needed.

These are the general operating requirements. To see how a nuclear power plant will fulfill them, it will be necessary to describe the plant, and then to relate it to the requirements.

Hypothetical Nuclear Power Plant

Basic Power Generation Scheme

The basic elements of a steam generator utilizing nuclear energy from a pressurized water-moderated and water-cooled reactor are shown in Fig. 1. The term water-moderated refers to the use of water as the medium in which the neutrons liberated in the process of fission are slowed down to an energy level at which they are effective in producing additional fissions. The uranium used is not receptive to high-energy neutrons. In the reactors described here, ordinary water is used for this function as well as to carry away the heat generated.

The core, in which the heat is liberated, is contained in a pressure vessel through which water is channeled. The water, having absorbed the heat of fission, passes out to the heat exchanger where the heat is transferred to the secondary water from which the steam is made. The cooler water leaves the heat exchanger and returns to the pressure vessel and core. Circulation is maintained by the circulating pump. To generate steam at temperatures greater than 212 degrees Fahrenheit, the entire primary system is under high pressure so that boiling does not occur in the core at the elevated temperatures at which the primary plant operates. This pressure is regulated in the pressurizer by a spray and heater control system to a reasonably constant value.

Of interest to the user of the steam is the fact that, although the water in the primary system becomes radioactive while it passes through the core, where it is subjected to intense neutron irradiation, its radioactivity is not in the form of neutron production so that it does not make the secondary, or steam side of the plant radioactive. Hence, all parts of the plant containing the primary system must be shielded, but the parts containing the steam system alone require no shielding.

Of further interest is the fact that while the plant is in operation all operations under the shield must be by remote control. Moreover, it is desirable to reduce the leakage from the primary system in order to reduce the amount of radioactive material that must be processed for disposal. The system described here is designed for zero leakage. Hence, the valves, pumps, and mechanisms for regulating the control rods must be of special design. The problems associated with this equipment and their solution have been described in other papers.[1,2]

CONTROL OF THE REACTOR

When the plant is at steady state, the reactor core will produce as much heat as the steam system removes from the water at the heat exchanger. A characteristic of reactors of the type described here is that they may be designed to have a negative temperature coefficient of reactivity. This means that if the temperature of the water in the core drops, the activity of the reactor will increase and

---

Paper 55-749, recommended by the AIEE Nucleonics and Power Generation Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted November 18, 1954; made available for printing July 29, 1955.

N. E. WILSON is with the Westinghouse Electric Corporation, Pittsburgh, Pa.

This paper was prepared under *Contract AT-11-2-GEN-14* with the Atomic Energy Commission.

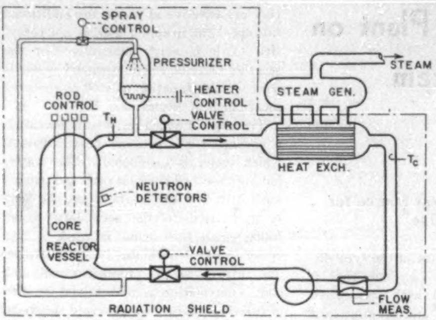


Fig. 1. Schematic diagram of nuclear power plant

so, therefore, will the power level and heat generation, thus raising the water temperature again. Conversely, a rise in water temperature will cause the activity and power level to decrease until the water temperature has returned to its stable value. This inherent feature results in an exceptionally stable plant, and at the same time enables it to be self-regulating on load changes. This can be visualized by a study of Fig. 2. In this diagram, the temperature of the water leaving the reactor T_H, the temperature of the water leaving the boiler T_C, the steam power load, and the reactor heat output are plotted as functions of time during a power change. It is assumed that steam power output is increased rapidly to a new level in less time than it takes for the water to travel from the boiler to the reactor. The reactor power does not change until the colder water reaches the core. The colder water causes the reactor power level to increase, increasing the temperature of the water leaving the reactor. The power level of the reactor will increase as long as the average temperature of the water is below normal. As soon as the reactor has restored the energy initially taken from the water, it will settle to the new heat output needed to maintain the power flow equilibrium. The action described occurs without external regulation. Its effect is to restore reactor (and boiler) average temperature to its former value following the transient caused by a load change.

During the production of energy by nuclear fission, various end products of the fission process accumulate within the fuel. These are the elements into which uranium splits. Some of them are initially formed as an isotope of one element and decay with time to another, and with further time to still another. One of these of interest is an isotope of xenon, Xe^{135}, which is important because it is highly effective in absorbing neutrons and thus preventing them from being useful in fission. The absorptive effect is similar to that of the control rods. This isotope of xenon is formed as a decay product of iodine which is substantially a fission product.

The Xe^{135} is eliminated by two means. It will decay radioactively to an isotope of cesium, but its decay takes longer than its formation. In the presence of a neutron flux, i.e., when the reactor is at power, it is changed by neutron absorption into Xe^{136}. Both the Cs^{135}, and the Xe^{136} are essentially nonabsorptive and hence produce no effect on the operation of the reactor. A more detailed discussion of these effects may be found in textbooks on reactor technology.[3,4] During any prolonged steady-state operation of the reactor, the amount of the Xe^{135} reaches an equilibrium value. Following any power change the quantity of this isotope changes over a period of time measured in hours. The effect of this is to alter the temperature at which the reactor is self-regulating, as described previously. To maintain the operating temperature constant the control rods must be adjusted. This adjustment may be manual or automatic, the automatic control being based on temperature.

The result of the control features just described is shown in Fig. 3. On this diagram, the various reactor temperatures are plotted as a function of plant power output for steady-state conditions and constant flow of the primary coolant. The average temperature of the reactor is maintained constant by regulation of the control rods as the amount of xenon changes. The reactor outlet temperature increases and the reactor inlet temperature decreases as the plant output increases. The effect of maintaining the average temperature of the primary coolant constant is to produce a steam temperature, and hence steam pressure, which varies with load. In a unit system, this variation of steam pressure, although not conventional, is not detrimental to plant operation.

If the plant is to be scheduled for light load, the coolant flow may be reduced by a reduction in pump speed. The result of this operation is illustrated by Fig. 4. The pumps are operated at one-half speed for loads to 50 per cent (%) of plant capability, and at full speed for loads from 50% to 100% of plant capability. The advantage of this is the considerable savings in pumping power for light-load operation. The power required for the pumps at one-half speed is only one-eighth the power required for full speed.

Operation of Nuclear Power Plant in Electric System

Plant Start-up

The reactor plant at ambient temperature, with the plant checked out prepara-

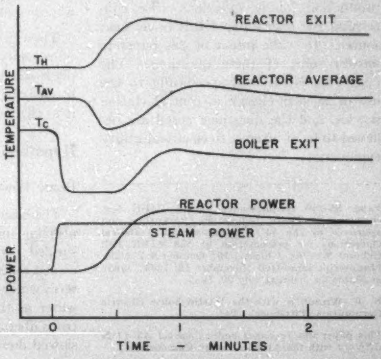


Fig. 2. Reactor power and primary plant temperatures following a change in steam power output

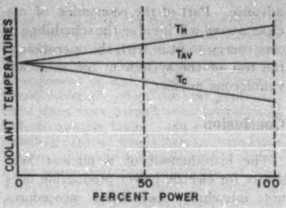


Fig. 3. Temperatures versus power. Pumps at full speed

tory to making the reactor critical, corresponds to the conventional coal or oil fired plant at the time of lighting off the burners. With the reactor plant ready for start-up, the pumps will be running at reduced speed. Usually a primary system pressure a little above ambient is maintained on the primary loops to prevent the corrosion inhibitor from coming out of solution. Starting with this condition, the general sequence of events is as follows:

1. The control rods are withdrawn at a controlled rate to approach criticality. The condition of the reactor can be determined by observation of the nuclear instruments which indicate the multiplication of the neutron population within the core. Approximately 20 minutes is required to attain criticality for a hypothetical central station reactor.

2. The reactor power, when criticality is reached, will be many powers of 10 below normal power level. By adjustment of rods the reactor will be placed on a "positive period," which means that power will multiply at a regular rate until normal power range is reached. It will take 10 to 15 minutes to raise power to the vicinity of 10-per-cent power where plant warm-up can begin.

3. With the reactor at a power level sufficient to raise temperature at 150 to 200 degrees Fahrenheit per hour, approximately 3 hours will be required to warm the plant to the operating temperatures of present-day technology. The primary loops are pressurized by operation of the heaters in the pressurizer to keep pace with the increasing temperature.

4. If desired, the turbine may be rolled as soon as adequate steam pressure is available. Load may be placed on the turbine as soon as it can be brought to operating speed.

5. When the primary plant reaches design temperature it may be placed in automatic control which will operate to hold the average temperature constant.

The process of starting up and of bringing the plant to full power operation thus requires $3^1/_2$ to 4 hours if it is initially at ambient temperature. If the plant is at or near design temperature, as it would be following an overnight shutdown, the time would be about 1 hour, consisting of the time required to attain criticality and power range operation plus the time required to place the turbine in operation and bring it to full load.

Plant Operation in Power Range

There is essentially no inherent restriction on the rate at which the reactor plant can pick up load. The rate at which the turbine can assume load is the determining factor. For decreasing loads, the rate at which load is removed will affect the design of the primary plant in that the pressurizer must be capable of keeping the primary pressure within design limits. The size of step load reductions that can be accommodated by a nuclear power plant of the type considered in this paper will depend on the size of the pressurizer and the volume of the primary system. Nevertheless, without oversizing the pressurizer above the size required for maintenance of normal primary water inventory, the reactor plant is able to accommodate reductions in load of 20% to 30% of the plant rating as fast as the turbine throttle can close for load changes, normally 2% to 3% per second. It is able to take a complete load cutback from 100% minimum or zero load at a rate of 5% to 10% per minute. These rates of load change more than meet the requirements of central station practice.

Emergency Load Reduction and Recovery

Two types of emergency load rejection are of interest; generator-breaker-trip, and steam throttle-trip. In either case the steam pressure relief valve will probably open. It is not necessary to shut down the reactor if the transient primary loop temperature rise can be prevented from exceeding a design limit by controlled steam exhaust until the reactor power has cut back. In this case the unit may be restored to service as fast as the steam and electric plant can be recovered.

If the reactor emergency shutdown is tripped during these emergencies, the rods will have to be withdrawn again, and criticality regained. The process of restoring the reactor to power level will take 15 to 30 minutes, depending on the mechanism for safety shutdown and the length of time before rod withdrawal is begun.

Plant Shutdown and Start-up After Normal Shutdown

The process of shutting down a reactor is no more involved than shutting down a conventional plant. After the steam load has been reduced to zero or minimum, the rods are inserted to make the reactor subcritical. However, two factors, decay heat and the xenon effect, are present, which make the nuclear power plant differ from the conventional. A discussion of these factors and their effect on plant operation follows.

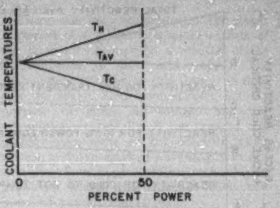


Fig. 4. Temperatures versus power. Pumps at half speed

Insertion of the rods causes the reactor to become subcritical and reduces the production of fission heat to a fraction of a per cent essentially at once. However, the radioactive decay of fission products within the core will produce heat equal to a few per cent of full power immediately following shutdown from full power operation. This decay heat power will decrease exponentially with time. There will be a period of several hours during which the decay heat is greater than can be dissipated by thermal radiation from the primary plant. Hence some circulation of primary coolant must continue through the core, and means must be provided for removal of the heat from the primary system. This may be in the form of a small auxiliary condenser fed from the steam mains.

The formation of the isotope Xe^{135} has been referred to. It was stated that during any steady power operation, the formation of Xe^{135} and its destruction by neutron flux reach an equilibrium state. When power is decreased to zero the destruction of Xe^{135} by neutron absorption ceases. Hence the amount of xenon present increases quite appreciably until its loss by radioactive decay becomes greater than its formation by radioactive decay of its parent isotope, I^{135}. The operating effect of xenon, following a shutdown, is to require greater withdrawal of control rods to attain criticality on restarting. Coincidentally, the time at which this effect is at a maximum following shutdown, is approximately the same as a normal overnight plant shutdown. Since the amount of xenon that will be present is such a complicated function of the operating history of the reactor, the position of the rods for criticality can never be known in advance.

ol. 74
. 1
955

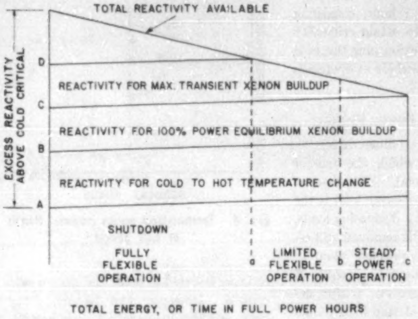


Fig. 5. Reactivity: energy diagram

Effect of Core Age on Plant Operation

When a particular core is new, it will have the ability to maintain criticality with any amount of xenon that may accumulate as the result of operating history. Operationally, this means that the plant is fully flexible with respect to the size of power changes to which it is subjected. It may be shut down after extended power operation and started up at any time thereafter. As the amount of energy taken from the core increases, the excess reactivity available decreases through the depletion of fissionable uranium, until a point is reached at which the peak xenon can no longer be overridden at design temperature. The equivalent full power hours of operation obtainable before this condition occurs is subject to design.

The condition described is illustrated by the diagram of Fig. 5, where the components of reactivity are plotted as ordinates, with total energy removal, or equivalent full power hours, as abscissa. Thus, when the core is new, it can be critical at *A* if it is cold. As the plant warms up, the rods must be withdrawn to *B* because of the negative temperature coefficient of reactivity. If the plant is brought to full power and operated there for several hours, equilibrium xenon for full power will be built up and the rods must be withdrawn to *C* to maintain criticality. If the power is reduced to zero, or very low value, transient xenon will build up, and in about 8 to 10 hours the rods would need to be out to *D* to maintain criticality or to restart the plant if it had been shut down.

From a study of Fig. 5 it can be seen that up to time *a*, in full power hours, there is always enough reactivity to enable the plant to be made critical and reach full power for any operating condition. Operation beyond point *a* is possible, and, depending on core design, there may be considerably more energy usable than was obtained in reaching point *a*. The plant may be operated at full power at full design temperature up to point *c*. Between *a* and *b* slightly limited operation may be obtained in that the plant can be subject to normal load swings, but if it should be shut down for 8 to 10 hours, it would be necessary to start up at reduced temperature. If the temperature at which criticality can be obtained is high enough so that an appreciable amount of power can be developed, say 50%, then the transient xenon will burn out sufficiently in 2 or 3 hours to permit operation at full temperature and power. Thus point *b* is partially determined by the design of the steam plant and turbine.

Beyond point *b*, the plant can be operated at or near full load as a base load plant until point *c* is reached. Beyond point *c* operation would require that plant temperatures be reduced an increasing amount as operation continued.

It is interesting to note that the life cycle of a particular core is the reverse of the whole life cycle of a conventional steam plant. Most steam plants are used as base load plants when new because their efficiency is higher than existing plants. As newer plants come on the line, the existing ones are used increasingly as swing or regulating plants. A new core in a nuclear energy plant has the ability to operate with full flexibility when it is new. As its age increases, some of the operational flexibility is lost, but much additional energy may be obtained without incurring additional fuel-handling charges for reloading by continuing its operation as an increasingly base-loaded plant.

Since the points *a*, *b*, and *c* are a function of design, operation of the plant during the life of a core can be scheduled in advance. Part of the economics of nuclear power will involve the scheduling of core energy to fit in with the operation of the rest of the system in the most advantageous manner.

Conclusion

The introduction of a nuclear heat source for electric power production does not introduce operational procedures that are markedly different from conventional practice. The operator of such a heat source will find new names on his instruments and controls, but when he has learned their meaning, he will find that he is using the plant in the old and familiar way.

References

1. Position Control in Sealed Systems, **W. H. Esselman, W. H. Hamilton.** *Proceedings* of the 1953 Conference on Nuclear Engineering, California Book Company, Ltd., Berkeley, Calif., 1953.

2. A Very Low Speed Drive with Flooded Rotor Motor, **W. E. McCown.** *Electrical Manufacturing*, New York, N. Y., vol. 54, Aug. 1954, pp. 141–43.

3. Elements of Nuclear Reactor Theory (book), **S. Glasstone, M. C. Edlund.** D. Van Nostrand Company, Inc., New York, N. Y., 1952.

4. Introduction to Nuclear Engineering (book), **R. L. Murray.** Prentice Hall, Inc., New York, N. Y., 1954.

◆

Discussion

Harold E. Gove (Union Electric Company of Missouri, St. Louis, Mo.): Utility engineers should find this well-written paper a valuable introduction to a number of basic operating problems associated with nuclear power.

An especially significant characteristic of the pressurized water reactor cited by the author is the dependence of operating flexibility on core age as measured by total energy extraction. With any particular core a point is reached beyond which the reactor is no longer adequately responsive to load swings nor capable of rapid startup after shutdown. Finally, as burnup of fuel continues still further there will come a time when only base-load operation is practicable with the depleted core.

This situation points up an advantage of the homogeneous reactor which utilizes a liquid fuel. For in this case a small fraction of the fuel may be continuously bled from the reactor and circulated through a reprocessing facility so that reactivity and hence operating flexibility can readily be maintained. Although the technology of the homogeneous reactor has not yet reached a stage where commercial projects have been initiated, it may in time become a strong competitor of present types because it offers the fundamental advantages of the continuous process as contrasted with the batch process.

In the case of the pressurized water reactor there is, first of all, the economic question as to whether or not the extra cost of maintaining full operating flexibility can be justified. Periodic partial reloading or rearrangement of fuel to hold up reactivity would adversely affect per-unit generating costs since more frequent shutdowns would reduce capacity factor; and fuel and fuel-handling expense would increase with lesser burnup. Obviously we will need to accumulate some operating experience at least with the "first generation" of commercial reactors yet to be built, before economic evaluations will have much substance. One can remark however, that nuclear plants will not be base-loaded if off-peak energy can be produced more cheaply with other system capacity. Estimates of nuclear power cost usually assume high-capacity-factor operation, implying relatively low marginal cost, but the latter may or may not be attained in early designs.

However, apart from the question of economics it is clear that in the hypothetical case of a power system wholly dependent on several nuclear reactors of the pressurized water type it would scarcely be feasible to allow any core to age beyond full flexibility. Even in a large system where there are only one or two *PWR* units many operating people would doubtless be reluctant to accept the idea of "crippled" units as a normal component of system capacity. Moreover in the nature of nuclear power designs conceived to date, the economic advantage of bigness is decisive. Thus the unavailability of a nuclear reactor unit is admissable only if unavailability of a sizeable block of capacity can be tolerated. Some compromise may be possible, for example, perhaps several more heavily enriched fuel rods could be held in reserve, so that in the event of an emergency it would be possible promptly to convert a reactor with depleted core capable only of base-load operations, to fully flexible operation.

N. E. Wilson: The author wishes to thank Mr. Gove for his discussion of the paper, and comments on the economic aspect of various types of reactors.

There was no comparison in the paper discussed of a pressurized water reactor of the fixed-core type with a reactor of the homogeneous or any other type. The author believes that at this time it is difficult to make such comparisons for the purpose of citing advantages and disadvantages—especially in the area of economics—since some of the reactors which appear to have marked advantages in certain areas have as yet unsolved technological problems which will influence their economics for some time to come.

The author agrees with Mr. Gove's comments in the last paragraph of the discussion to the effect that technical developments may conveivably be made which would remove some of the inflexibility present in a core of the fixed type as it approaches the end of burnup. The advancement of these developments and of the day when sufficient information is available for economic evaluations can only be enhanced by the programs under way for building and operating reactors of all types.

Presentation of Data on Broad-Band and Pulse Transformer Cores

P. R. GILLETTE
ASSOCIATE MEMBER AIEE

KAZU OSHIMA
NONMEMBER AIEE

K. W. HENDERSON
ASSOCIATE MEMBER AIEE

R. M. ROWE
NONMEMBER AIEE

l. 74
1
55

ADEQUATE information concerning the characteristics of available core materials is an essential tool of the transformer engineer. Not only must the information be accurate and complete but it should be presented in a form convenient for use. Present standard methods of presenting such information for use in designing pulse transformers are reviewed herein, and an improved method is proposed; this proposed method can also be employed, with slight modifications, for presenting data to be used in designing broad-band units. The new method is derived from an accurate equivalent circuit representation of a test core and coil, and enables the transformer engineer to predict the response and loss characteristics of a design with much greater accuracy and greater ease than is possible with data as now commonly presented.

It is well known that the most important core-dependent factors in the behavior of power transformers are the core loss and no-load exciting current; it is also well known that accurate predictions of these factors can be made with the aid of curves giving watts per pound and exciting volt-amperes per pound as a function of flux density at the power frequency. In broad-band and pulse transformers, on the other hand, the response characteristics are much more important than the no-load current as such. The prediction of transformer response characteristics always involves the use of an equivalent circuit representation of the transformer core (though in some types of calculation this fact may not be obvious), and the core data curves published for use in designing broad-band and pulse transformers are used to calculate the values of the elements in the equivalent circuit. The equivalent-circuit representation of a test core and coil may take the form of a simple inductance, Fig. 1(A), an inductance shunted by a resistance, Fig. 1(B), an inductance shunted by a series combination of a resistance and an inductance, Fig. 1(C), or any one of these three combinations in series with a resistance.

Two methods of obtaining and presenting data for use in designing broad-band transformers are in common use. In the first method, the permeability of the material, defined as the slope of the straight line between the two tips of a hysteresis (**B-H**) loop, is obtained either from measurements of **B-H** loops or from measurements of exciting volt-amperes made on a test core and coil; either procedure implies the assumption of an equivalent circuit consisting of a single inductance. The measurements are often made with various amounts of direct current in the coil to obtain values of incremental permeability. The resulting data are usually presented as curves of permeability versus peak flux density at various frequencies. Curves of core loss versus flux density at the same frequencies are also presented; these curves are based on the assumption that the equivalent circuit is as shown in Fig. 1(B). This method, which corresponds closely to the one used for power frequencies, is commonly used in the audio-frequency range. In the second method, an inductance measurement is made on a *Q*-meter or other type of bridge which gives the impedance of the test

Paper 55-750, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3-7, 1955. Manuscript submitted June 8, 1955; made available for printing July 27, 1955.

P. R. Gillette, Kazu Oshima, K. W. Henderson, and R. M. Rowe are with the Stanford Research Institute, Menlo Park, Calif.

The work described herein was carried out at Stanford Research Institute under *Contract No. DA 36-039 SC-42645* with the Signal Corps Engineering Laboratories. The authors wish to thank the cognizant representatives of those laboratories for permission to publish results of the work in this paper.

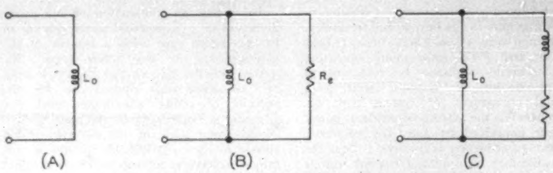


Fig. 1. Common types of equivalent circuit for transformer core

core and coil in terms of an inductance in series with a resistance; the effect of the coil can be eliminated by making measurements on it both with and without the core in place. Since the exciting voltage is usually small, the permeability obtained in this manner is the initial permeability of the material. The resulting data are normally presented as curves of permeability versus frequency and Q versus frequency. This method, originally developed for air-core coils, in which the loss is all in the winding and thus is accurately represented by a series resistance, is commonly used at radio frequencies.

The method commonly employed for obtaining and presenting data for use in designing pulse transformers was established by a committee of engineers representing core manufacturers and transformer manufacturers, sponsored by the Massachusetts Institute of Technology Radiation Laboratory.[1] In this method, **B-H** loops are obtained at flux density swings ranging up to 10,000 gausses, at four different values of pulse duration—0.25, 0.5, 1.0, and 2.0 microseconds. One set of curves consists of the loci of the type of these **B-H** loops, one locus for each pulse duration. The other set consists of plots of the areas of these loops, in ergs per cubic centimeter per cycle, as functions of the flux density swing for the four pulse durations. The first set of curves can easily be converted into curves of permeability versus flux density swing for the various pulse durations. The curves thus obtained correspond exactly to those presented for audio transformer core materials, and imply the use of the same types of equivalent circuit—a simple inductance for the permeability curves, and an inductance in parallel with a resistance for the loss curves.

The chief objection to the methods of measurement and presentation in common use for both broad-band and pulse transformers is that the effective values of the elements in the equivalent circuits on which they are based vary markedly with frequency or pulse duration. When an equivalent circuit whose elements vary in this manner is used in calculating frequency response, new circuit element values must be used in calculating each point on the curve. Similarly, when calculating output pulse top shape, new element values must be used in calculating each point along the top of the pulse. It would obviously be more convenient to use an equivalent circuit with constant element values in such calculations. It will be shown in the following that, for most core materials, a circuit of the form shown in Fig. 1(C) will have constant-valued elements, and core data can be presented so that the element values can easily be calculated for a given transformer design.

Theory

Theoretical analyses presented by the M.I.T. Radiation Laboratory[2] suggest the possibility of representing a pulse transformer core and test coil by an equivalent circuit composed of constant-valued elements. It is shown that a silicon steel core can be represented, for flux densities up to the value at which the lamination surfaces begin to saturate, by an equivalent circuit consisting of an infinite number of shunt branches. The first branch is an inductance L_0, and each of the others is an inductance and a resistance in series. The values of these elements are given by the following expressions, in which n is the number of the shunt branch

$$L_0=\frac{4\pi\mu}{10^9}\left(\frac{N^2A}{l}\right)\text{ henrys} \tag{1}$$

$$L_n=\frac{L_0}{2}\text{ henrys}\quad(n=1, 2, 3, \ldots) \tag{2}$$

$$R_n=\frac{2\pi^2\rho n^2}{d^2}\left(\frac{N^2A}{l}\right)\text{ ohms}\quad(n=1, 2, 3, \ldots) \tag{3}$$

where

μ = instantaneous d-c incremental permeability of the material
ρ = resistivity of the material, ohm-centimeters
d = lamination thickness, centimeters
A = cross-sectional area of the core, square centimeters
l = mean magnetic path length of the core, centimeters
N = number of turns in the winding

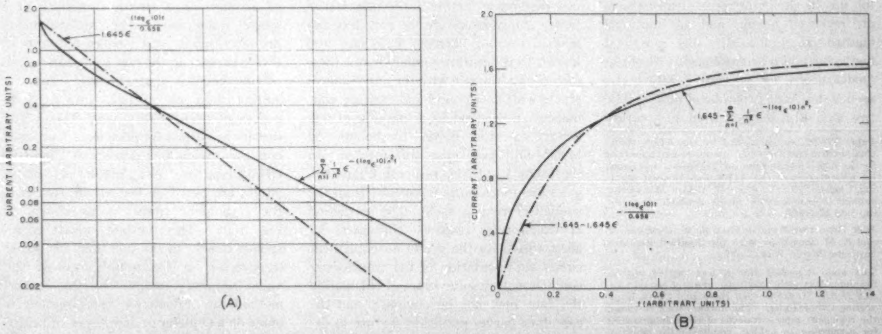


Fig. 2. Comparison of curves for L_eR_e branch and for infinite number of L_nR_n branches

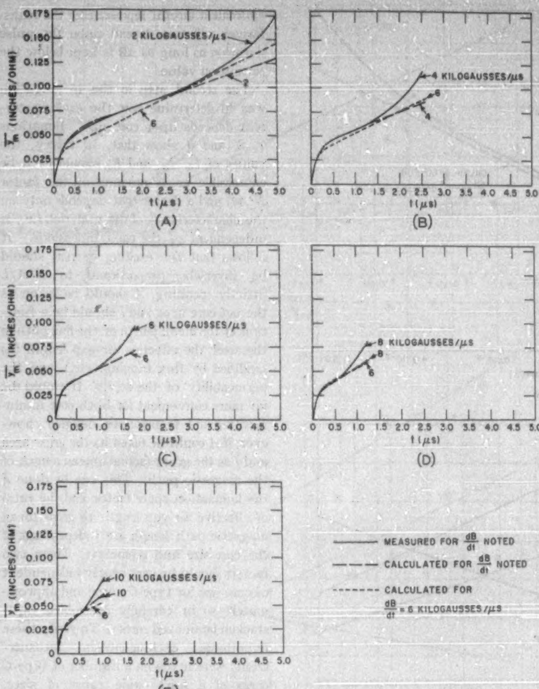


Fig. 3. Comparison of calculated and measured $\bar{y}_m$ curves

ol. 74
. 1
955

The current flowing in the first of these branches under application of a step-function voltage of amplitude E_0 is

$$i_0=\frac{E_0t}{L_0} \tag{4}$$

and the sum of the currents in the remaining branches is

$$i=\sum_{n=1}^{\infty}\frac{E_0}{R_n}\left(1-\epsilon^{-\frac{R_n}{L_n}t}\right) \tag{5}$$

The value of this sum has been calculated as a function of time on the assumption that $\sum_n(E_0/R_n)=1$, and $R_1/L_1=\log^e10$; this means that the exponential term for $n=1$ falls to 1/10th of its initial value in unit time. The resulting curve is shown in Fig. 2.

An equivalent circuit with an infinite number of elements cannot be used as a basis for a practical design procedure, so consideration was given to the possibility of approximating the complete circuit by a circuit of two branches, one branch containing the inductance L_0 and the other containing an inductance L_e and a resistance R_e in series; see Fig. 1(C). The approximate circuit should give the same final current amplitude and the same value for energy flow into the core during a long pulse as that given by the complete circuit. Hence, the following relations must be satisfied by L_e and R_e; see reference 3.

$$\frac{1}{R_e}=\sum_{n=1}^{\infty}\frac{1}{R_n} \tag{6}$$

and

$$\int_0^{\infty}\epsilon^{-\frac{R_e}{L_e}t}\,dt=\sum_{n=1}^{\infty}\int_0^{\infty}\epsilon^{-\frac{R_n}{L_n}t}\,dt \tag{7}$$

The values of R_e and L_e which satisfy equations 6 and 7 are

$$R_e=\frac{12\rho}{d^2}\left(\frac{N^2A}{l}\right) \tag{8}$$

$$L_e=\frac{L_0}{5} \tag{9}$$

The current in the first branch of this simplified circuit is the same as that in the first branch of the complete circuit; the current in the second branch is

$$i_e=\frac{E_0}{R_e}\left(1-\epsilon^{-\frac{R_e}{L_e}t}\right) \tag{10}$$

If $R_1/L_1=\log^e10$, then $R_e/L_e=\log^e$10/0.658. The value of i_e has been calculated as a function of time, with this value of R_e/L_e, and with $E_0/R_e=1$, and the resulting curve has been plotted in Fig. 2 for comparison with the corresponding curve for the complete circuit. Since the errors involved in practical procedures for measuring pulse exciting current amount to at least 2 per cent and are greatest near the beginning of the pulse, and since variations among cores in a given production lot are usually of the order of 10 per cent, the 3-element circuit should be sufficiently accurate for most practical purposes.

Some of the assumptions on which the foregoing theory is based are of questionable validity. In addition, a number of other equally questionable assumptions are involved in the application of the theory to cores as used in transformers. In type-*C* cores which are formed by winding thin magnetic-strip material on a mandrel, annealing the resulting core, impregnating it with a plastic binder, and then sawing it into two C-shaped halves, the thickness and conductivity of the interlaminar insulation, the shape and dimensions of the air gap, and the quality of the machining and etching of the cut surfaces are important factors which have not been taken into account in the theory. The effects of these factors, and to a lesser degree the intrinsic characteristics of the material, can vary over considerable ranges among cores of the same size and production lot. An experimental investigation was therefore undertaken to determine whether or not the behavior of transformer cores actually can be represented by a 3-element equivalent circuit and, further, what factors produce significant variations in the element values.[4]

Experimental Investigation

Most of the measurements discussed were made on type-*C* cores of 2-mil grain-oriented silicon steel. The conclusions resulting from analysis of the data obtained in these measurements were checked by making a few measurements on stacked cores, of 6-mil 48-per-cent nickel alloy and on ferrite cores. The voltage source used in making pulse measurements was a hard-tube pulse

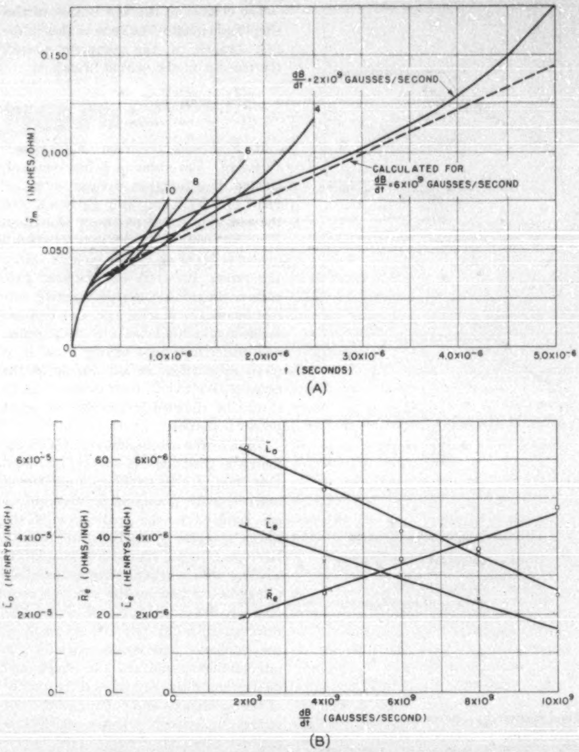


Fig. 4. Typical characteristics of type-C 0.002-inch oriented silicon-steel cores, with no gap spacers added

generator, with four type *4PR60A* tetrodes in parallel as the output stage, and with a 1,000-ohm noninductive resistor connected across the output terminals in parallel with the test core and coil assembly. Current pulse shapes were observed with a Tektronix type *513D* oscilloscope using noninductive current-measuring resistors.

The first step in the investigation was to study the dependence of the exciting current pulse shape of a given core on the flux density swing $\Delta\mathbf{B}$ or the pulse duration T at a given value of $d\mathbf{B}/dt$, i.e., a given value of applied voltage. Since a core has no way of "knowing" at the beginning of a pulse how long the pulse is going to be, the only conditions which can affect the rate at which exciting current builds up during the pulse are the magnitude of the applied voltage and the initial flux density in the core. If there is an air gap in the core, variations in initial flux density caused by variations in reverse current between pulses will usually be fairly small. Furthermore, the most important effect of a change in initial flux density will be a change in the value of $\Delta\mathbf{B}$ at which the core begins to saturate. Hence, if the pulse voltage is kept constant and the pulse duration is varied, with the product of voltage and duration kept less than enough to saturate the core, the exciting current should always be the same function of time, the curves for long pulses simply being extensions of those for shorter pulses. Appropriate exciting-current measurements have been made to verify this conclusion, and it has been found to be valid. A corollary of this conclusion is that, for a given value of $d\mathbf{B}/dt$, the values of the elements in an equivalent circuit representing the transformer will not depend upon the pulse duration as long as $\Delta\mathbf{B}$ is kept below the saturation value.

The second step in the investigation was to determine how the exciting current depends upon core size. Equations 1, 8, and 9 show that, in theory, the values of L_0, L_e, and R_e should each be expressible as a product of the factor N^2A/l and a factor that depends only on the characteristics of the material, i.e., is independent of the core dimensions. It follows that the exciting current should be inversely proportional to N^2A/l. Strictly speaking, A should be taken as the net core area and l should be a function of the mean length of the flux path in the steel, the effective air-gap length (as modified by flux fringing, etc.), and the permeability of the steel. It would be far more convenient for both core manufacturer and transformer designer, however, if A could be taken as the gross area and l as the gross (actual) mean length of the magnetic path. This can be done if the lamination space factor and the ratio of effective air-gap length to gross mean magnetic path length are independent of the core size and geometry. These two factors should be very nearly independent of core size for type-*C* cores, and approximately so in carefully assembled butt-stacked laminated cores. To verify these hypotheses, exciting-current measurements were made on a number of type-*C* cores of a fairly wide range of sizes. Each current value was multiplied by N^2A/l, where A was the gross core area and l the gross mean magnetic path length of the core, as calculated from the nominal dimensions listed in the core catalogue. When the resulting "normalized" curves were compared, it was found that the variations among cores of the same size were just as great as the variations among cores of different sizes.

The third step in the investigation was to determine whether or not an equivalent circuit composed of three elements L_0, L_e, and R_e will give an exciting-current pulse of the same shape as the current pulse actually observed. In this part of the investigation, exciting-current pulses were obtained for five values of $d\mathbf{B}/dt$. For convenience in analyzing these curves, new variables $\bar{y}_m$, $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ were defined. The quantity $\bar{y}_m$ is the exciting current i_m multiplied by N^2A/E_0l, or

$$\bar{y}_m = i_m \frac{N^2A}{E_0 l} \qquad (11)$$

The other three variables are defined by the relations

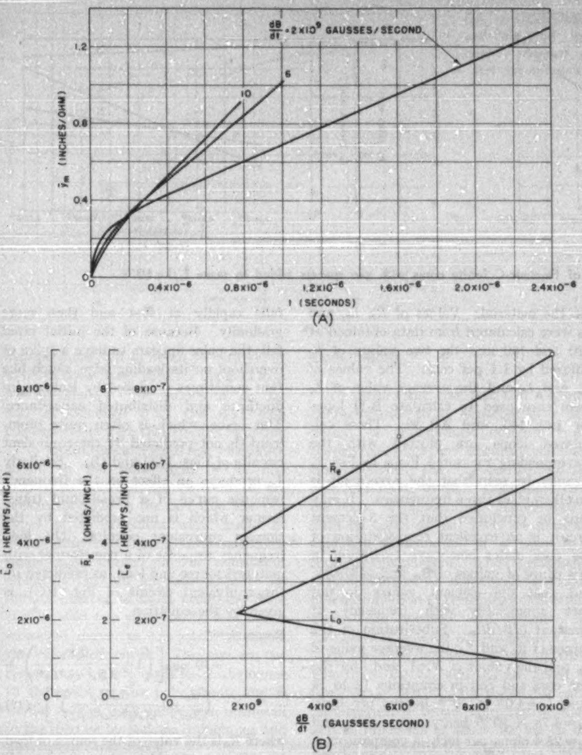


Fig. 5. Typical characteristics of 0.006-inch 45- to 50-per-cent nickel-alloy laminations, butt-stacked

$$L_0 = \bar{L}_0 \frac{N^2A}{l} \qquad (12)$$

$$L_e = \bar{L}_e \frac{N^2A}{l} \qquad (13)$$

$$R_e = \bar{R}_e \frac{N^2A}{l} \qquad (14)$$

Curves of $\bar{y}_m$ versus t were derived from the exciting current pulses for the various values of $d\mathbf{B}/dt$ and each of these five curves was analyzed by the method described in Appendix I to obtain five sets of values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$. Each of these sets of values was then used to obtain a calculated curve of $\bar{y}_m$ versus t, with the aid of the following equation

$$\bar{y}_m = \frac{t}{\bar{L}_0} + \frac{1}{\bar{R}_e}\left(1 - e^{-\frac{\bar{R}_e}{\bar{L}_e}t}\right) \qquad (15)$$

Each of the five calculated $\bar{y}_m$ curves is plotted with the $\bar{y}_m$ curve derived from the corresponding observed exciting-current pulse in Fig. 3. (A calculated $\bar{y}_m$ curve for $d\mathbf{B}/dt = 6$ kilogausses per microsecond is plotted in each part of Fig. 3 and also in Fig. 4(A) to indicate the amount of error that would result if the values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ obtained at 6 kilogausses per microsecond were used for other values of $d\mathbf{B}/dt$.) In each instance, the two curves are in excellent agreement up to a value of t for which $\Delta\mathbf{B}$ is equal to about 6,000 gausses; at this point the observed curve begins to rise more rapidly indicating that the lamination surfaces begin to saturate at this value of $\Delta\mathbf{B}$. Hence the 3-element equivalent circuit is an excellent representation of the core and test coil up to 6,000 gausses. The representation becomes less accurate as $\Delta\mathbf{B}$ is increased above this point, but is reasonably good up to 8,000 gausses.

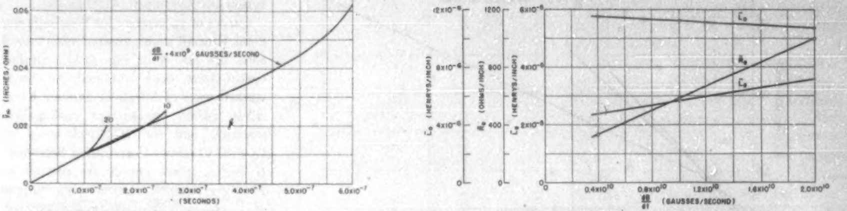
Curves of $\bar{y}_m$ versus t for Westinghouse 2-mil Hipersil grain-oriented silicon steel cores are shown in Fig. 4(A). Similar curves have been obtained for butt-stacked cores of Allegheny-Ludlum 6-mil *AEM* (*4750*) 48-per-cent nickel-alloy material, and for E-shaped cores of General Ceramics Corporation Ferramic-*C* ferrite; these curves are plotted in Figs. 5(A) and 6(A) respectively. Values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ have been calculated from each of the curves in these three figures; these values are plotted as functions of $d\mathbf{B}/dt$ in Figs. 4(B), 5(B), and 6(B). It may be seen that, for each type of core material, $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ are approximately linear functions of $d\mathbf{B}/dt$.

The fourth step in the investigation was to determine whether or not an equivalent circuit of this form will represent the sine-wave behavior of a core. Under application of a sine-wave voltage to a core and coil the value of $d\mathbf{B}/dt$ will be a sine function of time, rather than a constant as it is during application of a step-function voltage. For this reason there would presumably not be a direct correspondence between the values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ obtained for a pulse voltage of a given amplitude and values obtained for any sine-wave voltage. However, it seems reasonable to suppose that, since the pulse values of $\bar{L}_0$, $\bar{L}_e$ and $\bar{R}_e$ for a given material depend only upon $d\mathbf{B}/dt$, the sine-wave values would depend only on the amplitude of the sine function representing the value of $d\mathbf{B}/dt$ under sine-wave excitation. Since this amplitude depends only on the amplitude of the voltage applied to a test core and coil and not on the frequency of the applied voltage, it would be expected that the equivalent circuit element values would be independent of frequency.

The validity of these conclusions can be determined by making **B-H** loop measurements at constant voltage and variable frequency on a test core and coil. It can be shown that, if the equivalent circuit is of the form under consideration, the **B-H** loops can be represented by the following equation

$$H(t) = \frac{4\pi}{10^9}\left(\frac{N^2A}{l}\right)\left[\left(\frac{1}{L_0} + \frac{\omega^2 L_e}{R_e^2 + \omega^2 L_e^2}\right)\mathbf{B}(t) + \left(\frac{\omega R_e}{R_e^2 + \omega^2 L_e^2}\right)\sqrt{\mathbf{B}_m^2 - \mathbf{B}^2(t)}\right] \qquad (16)$$

This is an equation for an ellipse. If the observed **B-H** loops are of this form, values for L_0, L_e, and R_e can be obtained by determining the values of **B** at **H** = 0 and **H** at **B** = 0 from loops measured at two different frequencies, and making an



(A) (B)

Fig. 6. Typical characteristics of Ferramic-C ferrite cores with gap spacers added to make $l_g/l = 10^{-3}$

appropriate calculation. The calculation procedure is given in Appendix II; it involves a calculation of the values of L_e and R_e from the two values of **H** at **B**$=0$, and then calculating L_0 from the two values of L_e and R_e and one of the values of **B** at **H**$=0$. A comparison of these values of L_0 that can be calculated in this manner serves a check on the validity of the equivalent circuits. A further check is obtained by substituting the resulting values of R_e and L_e and an average of the two values of L_0 in equation 16, calculating **B**-**H** loops for the two frequencies at which the measurements were made, and comparing the calculated and observed loops. To determine whether or not the element values are independent of frequency, a **B**-**H** loop for a third frequency is calculated with the same equation and compared with an observed loop for the third frequency.

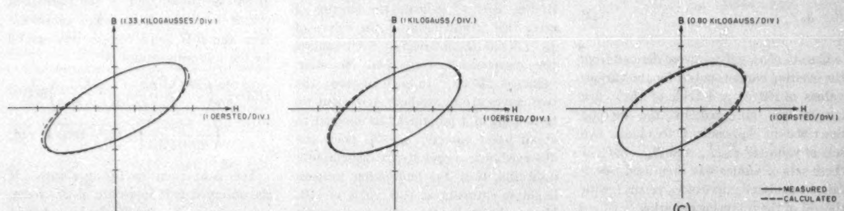
This process was carried out for a 2-mil Hipersil core. **B**-**H** loops were obtained at 120, 160, and 200 kc, and at the voltage required to make the peak value of $d\mathbf{B}/dt$ equal to 2.22×10^9 gausses per second (or slightly above the lowest value of $d\mathbf{B}/dt$ at which pulse measurements were made); the values of $\mathbf{B}_m$ corresponding to this value of $d\mathbf{B}/dt$ and the three frequencies are 2,960, 2,220, and 1,780 gausses respectively, all of which are far below the saturation value for the material. Values of R_e, L_e, and L_0 were calculated from data obtained at 120 and 160 kc; the two values of L_0 differed by 1.1 per cent. The values of R_e and L_e and the average value of L_0 were then used to calculate **B**-**H** loops for 120, 160, and 200 kc. These calculated loops are plotted with the corresponding measured loops in Fig. 7. It may be seen that the agreement is excellent at all three frequencies. It may thus be concluded that the 3-element circuit is an excellent representation of the core under sine-wave excitation in this range of values of $\mathbf{B}_m$ and $(d\mathbf{B}/dt)_m$, and that the element values do not vary appreciably with frequency at constant $(d\mathbf{B}/dt)_m$. Substitution of the values of R_e and L_e, the average value of L_0, and the values of N, A, and l for the test core and coil in equations 12 to 14 gave $\bar{L}_0 = 4.05 \times 10^{-5}$ henrys per inch, $\bar{L}_e = 4.78 \times 10^{-6}$ henrys per inch, and $\bar{R}_e = 25.4$ ohms per inch; a comparison of these values with the curves of Fig. 4(B) indicates that there is no simple relation between the sine-wave and pulse characteristics of this core material.

As has been indicated, the rate of rise of pulse-magnetizing current in a transformer is very high at the beginning of the pulse, but rapidly decreases to a much lower, constant rate determined by L_0. This results in an output pulse top which falls rapidly at first and then more gradually. Because of the initial rapid fall, the pulse appears to have a spike or overshoot on its leading edge, much like that sometimes produced by leakage inductance and distributed capacitance. This spike, which is often quite prominent, is not predicted by the equivalent circuits of Figs. 1(A) and (B). Similarly L_e produces an effect on the frequency response curve of a broad-band transformer which is not predicted by the simpler equivalent circuits. The low-frequency response of a transformer with matched source and load, as predicted by the equivalent circuit of Fig. 1(C), is given by the equation

Response =

$$-10 \log_{10}\left[\left(1+\frac{R_0R_e}{2(R_e^2+\omega^2L_e^2)}\right)^2+\left(\frac{R_0}{2\omega L_0}+\frac{R_0\omega L_e}{2(R_e^2+\omega^2L_e^2)}\right)^2\right] \quad (17)$$

where R_0 is the value of the source or load impedance, referred to the same impedance level as the other quantities. A frequency response measurement was made on the test core and coil on which the other sine-wave tests were performed, considered as a "1-to-1 autotransformer," with the result shown in Fig. 8. (It was impractical to use the radio-frequency amplifier with which **B**-**H** loops were obtained, so the frequency response



(A) f = 120 kc (B) f = 160 kc (C) f = 200 kc

Fig. 7. Calculated and measured B-H loops

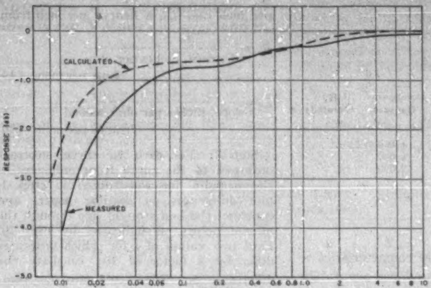


Fig. 8. Calculated and measured frequency response of test core and coil

measurement had to be made with about 1/100th the voltage used for the **B**-**H** loop measurements.) The values of L_0, L_e, and R_e obtained from the **B**-**H** loop measurements were then used in equation 17 to obtain the calculated response curve shown in the same figure. The discrepancies between the two curves are presumably due to the fact that the two types of measurement were made at widely different voltages, and because the single L_eR_e branch in the equivalent circuit is only an approximation to an infinite number of L_nR_n branches; note that at least two steps can be seen in the measured curve, as compared to one in the calculated curve. The rise in the response at high frequencies is analogous to the spike on the pulse front, and is similar to the peak sometimes produced in the reponse by leakage inductance and distributed capacitance.

Conclusions

To predict the core loss and response characteristics to be expected in a new design, a transformer designer must possess complete and accurate information concerning the characteristics of the core he intends to use. Commonly used methods of presenting representative data on cores do not give the designer complete information in a convenient form for this purpose. The investigation described has resulted in the development of a new method of presenting data which gives more accurate and complete information in a more convenient form.

In the new method, values of three quantities $\overline{L}_0$, $\overline{L}_e$, and $\overline{R}_e$ are given as functions of $(d\mathbf{B}/dt)_m$ for use in designing high-frequency broad-band transformers and as functions of $d\mathbf{B}/dt$ for use in designing short-pulse transformers. Values of the three equivalent circuit elements required to represent the core loss and low-frequency or pulse-top response characteristics of a transformer can be obtained from values of these three quantities by multiplying each quantity by the factor N^2A/l.

These curves are valid for a given type of material used in a given manner; e.g., one set of curves is applicable to Westinghouse 2-mil Hipersil type-*C* cores with no added gap spacers, and another is applicable to Allegheny-Ludlum 6-mil *AEM* laminations butt-stacked with no added gap spacers. For pulse transformers this lack of generality is an advantage rather than a disadvantage, except where core-resetting currents between pulses are important, since there will usually be one optimum ratio of air-gap length to total magnetic path length for each basic material, and the core stacking factor will be fairly constant. For broad-band transformers in which variable amounts of direct current may flow, it may be advisable to present the data in more general form and to incorporate the core-stacking factor and the ratio of air-gap length to total magnetic path length in the factors by which the quantities $\overline{L}_0$, $\overline{L}_e$, and $\overline{R}_e$ are multiplied to obtain L_0, L_e, and R_e.

Furthermore, the curves are valid only for flux densities within certain limits. The upper limit is set by core saturation. For pulse transformers, this limit may be indicated by publishing $\overline{y}_m$ versus t curves for various values of $d\mathbf{B}/dt$ $(\overline{y}_m = i_mN^2)$ A/E_0l. The rate of rise of these curves is very high for small values of t, and then decreases to a value which is constant for

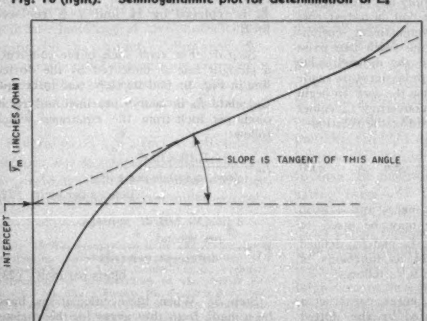


Fig. 9 (below). Determination of slope and intercept of $\overline{y}_m$ curve

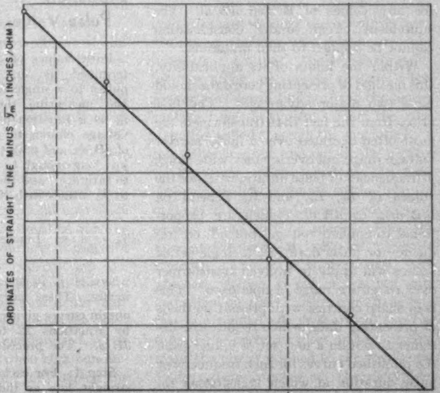


Fig. 10 (right). Semilogarithmic plot for determination of L_e

Table I

dB/dt, (Gausses per Second)	T, Microseconds							
	0.25		0.5		1.0		2.0	
	B, Gausses	H, Oersteds	B, Gausses	H, Oersteds	B, Gausses	H, Oersteds	B, Gausses	H, Oersteds
2×10^9	500	....	1,000	....	2,000	....	4,000	......
4	1,000	....	2,000	....	4,000	....	8,000	......
6	1,500	....	3,000	....	6,000	....	12,000	......
8	2,000	....	4,000	....	8,000			
10	2,500	....	5,000	....	10,000			
12	3,000	....	6,000	....	12,000			

a considerable length of time, as predicted by the equivalent circuit. For large values of t, the rate increases again in a manner not predicted by the equivalent circuit; the value of $\Delta\mathbf{B}$ corresponding to the value of t at which this discrepancy becomes excessive is the upper limit of applicability of the $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ curves. This limit is nearly independent of $d\mathbf{B}/dt$, however; hence, for most purposes it is sufficient to present the $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ curves and a single limiting value for $\Delta\mathbf{B}$. For sine-wave transformers, the upper limit on **B** is the value at which the **B-H** loop begins to show excessive departures from an elliptical form. This limit can be shown by publishing a family of **B-H** loops for a representative value of $(d\mathbf{B}/dt)_m$, since it will presumably be independent of $(d\mathbf{B}/dt)_m$. Here again, a specification of the limit itself will be adequate for most purposes.

In addition to this upper limit, there is apparently a lower limit on $\Delta\mathbf{B}$ or $\mathbf{B}_m$ below which the equivalent circuit becomes too inaccurate. This matter has not yet been completely investigated. It should also be pointed out that **B-H** loops for some materials may not be elliptical for any values of **B** and $d\mathbf{B}/dt$. The equivalent circuit under consideration cannot be applied to such materials.

Within the limits of its applicability, this method of presenting core data should have two major advantages. The first arises from the fact that transformers are most often operated over a fairly narrow voltage range but over a fairly wide range of frequencies or pulse durations. Since the values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ depend on $(d\mathbf{B}/dt)_m$ or $d\mathbf{B}/dt$, (which are proportional to applied voltage) but not on frequency or pulse duration, a single set of values will apply to a given transformer over its entire range of operation. This is in sharp contrast with present methods of presenting data, which require the designer to obtain a new set of values from the published curves for each frequency or pulse duration at which transformer response calculations are to be made.

The second advantage is that certain features of the response characteristics of a transformer can be predicted accurately with an equivalent circuit calculated from values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$, which cannot be predicted as accurately or at all with the type of data now being published. For example, the response of a broad-band transformer may be down about 1 decibel because of the effect of the L_eR_e branch of the equivalent circuit over a large portion of what would normally be considered the passband. This effect is not predictable from data as now published. Similarly, in pulse transformers a peak may appear at the beginning of the pulse top, the presence of which cannot be predicted from currently available pulse data. For these reasons it is believed that this proposed method of presentation should be investigated by core manufacturers and transformer engineers, and used as a basis for discussion in setting up new standards for test procedures and methods of presenting data on broad-band and pulse transformer core materials.

Appendix I. Determination of Pulse Values of L_0, L_e, and R_e

Pulse values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ are determined by first applying rectangular pulses to a single-layer coil on a test core and measuring the magnetizing current i_m as a function of time (with the pulse voltage chosen to give the desired value of $d\mathbf{B}/dt$, and with the pulse duration made just long enough to cause the core to begin to saturate), and then converting i_m values to $\mathscr{I}_m$ values with the aid of the relation

$$\mathscr{I}_m=\frac{i_m}{E_0}\left(\frac{N^2A}{l}\right)\text{ inches per ohm} \tag{18}$$

where A is in square inches and l is in inches. These curves may be used to obtain curves giving $\bar{R}_e$, $\bar{L}_0$, and $\bar{L}_e$, defined by equations 12 to 14, as functions of $d\mathbf{B}/dt$. The procedure is as follows:

Step 1: For each $\mathscr{I}_m$ curve, construct a straight line as indicated by the dotted line in Fig. 9. Find the slope and intercept of this line, and then calculate $\bar{R}_e$ in ohms per inch and $\bar{L}_0$ in henrys per inch from:

$$\bar{R}_e=\frac{1}{\text{intercept, inches per ohm}}\text{ ohms per inch} \tag{19}$$

$$\bar{L}_0=\frac{1}{\text{slope, inches per ohm-second}}\text{ henrys per inch} \tag{20}$$

Step 2: For each $\mathscr{I}_m$ curve, subtract ordinates of the curve from ordinates of the straight line constructed in step 1, plot differences on semilog paper, and construct the best straight line through the plotted points, as indicated by Fig. 10. Find two values of t for which ordinates differ by a factor of 10; calculate the difference Δt. Then calculate $\bar{L}_e$ from:

$$\bar{L}_e=\frac{\bar{R}_e\Delta t}{2.30}\text{ henrys per inch} \tag{21}$$

Step 3: When these calculations have been made from the curves for various values of $d\mathbf{B}/dt$, plot $\bar{R}_e$, $\bar{L}_0$, and $\bar{L}_e$ as functions of $d\mathbf{B}/dt$. In general, the best straight line through the values of each quantity will be a reasonable representation of the results, within experimental error.

Typical curves obtained in this manner for three classes of core materials are presented in Figs. 4 through 6.

For greatest accuracy, curves of these kinds should be obtained from measured magnetizing current curves. However, it is possible to obtain $\bar{R}_e$ and $\bar{L}_0$ curves from pulsed-excitation curves of the standard form specified in reference 1 as now published by many core manufacturers, provided the published curves are drawn with sufficient accuracy. The procedure is as follows:

Step 1: Set up a table of the form of Table I.

Step 2: Read the values of **H** corresponding to the various values of **B** listed for each value of pulse duration in Table I from pulsed excitation curves. Record these values in the table.

Step 3: Plot **H**-versus-**B** curves for the various values of $d\mathbf{B}/dt$ from the corresponding rows of data in the table. These curves will be as shown in Fig. 9, where $\mathscr{I}_m$ is replaced by **H** and t is replaced by **B**.

Step 4: For each such curve construct a straight line as indicated by the dotted line in Fig. 9, find its slope and intercept. Calculate $\bar{L}_0$ in henrys per inch and $\bar{R}_e$ in ohms per inch from the equations which follow:

$$\bar{L}_0=\frac{3.19\times10^{-8}}{\text{slope, oersteds per gauss}}\text{ henrys per inch} \tag{22}$$

$$\bar{R}_e=\frac{3.19\times10^{-8}d\mathbf{B}/dt,\text{ gausses per second}}{\text{intercept, oersteds}}\text{ ohms per inch} \tag{23}$$

Step 5: When these calculations have been made from the curves for the various values of $d\mathbf{B}/dt$, plot $\bar{L}_0$ and $\bar{R}_e$ as functions of $d\mathbf{B}/dt$.

Appendix II. Determination of Sine-Wave Values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$

Sine-wave values of $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ are obtained by first applying sine-wave voltages to a single-layer coil on a test core, and observing the resulting **B-H** loops on an oscilloscope by the usual method. Voltage amplitudes are chosen to give the desired values of $(d\mathbf{B}/dt)_m$ and measurements are made at two frequencies at each voltage. These loops may be used to obtain curves giving $\bar{R}_e$, $\bar{L}_e$, and $\bar{L}_0$ as functions of $(d\mathbf{B}/dt)_m$. The following quantities will be referred to in describing the procedure.

ω_a, ω_b = two angular frequencies, radians per second, at which measurements are made with a given applied voltage $(\omega_a<\omega_b)$.

$\mathbf{H}_a$, $\mathbf{H}_b$ = positive values of $\mathbf{H}(t)$, oersteds, at $\mathbf{B}(t)=0$ for **B-H** loops obtained at ω_a and ω_b respectively.

$\mathbf{B}_a$, $\mathbf{B}_b$ = positive values of $\mathbf{B}(t)$, gausses, at $\mathbf{H}(t)=0$ for **B-H** loops obtained at ω_a and ω_b respectively.

$\mathbf{B}_{ma}$, $\mathbf{B}_{mb}$ = peak positive values of $\mathbf{B}(t)$ for **B-H** loops obtained at ω_a and ω_b respectively.

$K_{\omega\mathbf{B}}=3.19\times10^{-8}\omega_a\mathbf{B}_{ma}=3.19\times10^{-8}\omega_b\mathbf{B}_{mb}$.

The values of these quantities are determined for the two **B-H** loops observed at a given value of $(d\mathbf{B}/dt)_m$, and the corresponding values of $\bar{R}_e$, $\bar{L}_e$, and $\bar{L}_0$ are obtained in the following manner:

Step 1: Calculate $\bar{R}_e$ from the equation

$$\bar{R}_e=\frac{K_{\omega\mathbf{B}}(\omega_b^2\mathbf{H}_b-\omega_a^2\mathbf{H}_a)}{\mathbf{H}_a\mathbf{H}_b(\omega_b^2-\omega_a^2)} \text{ ohms per inch} \qquad (24)$$

Step 2: Calculate $\bar{L}_e$ from either of the equations

$$\bar{L}_e=\frac{1}{\omega_a}\sqrt{\bar{R}_e\left(\frac{K_{\omega\mathbf{B}}}{\mathbf{H}_a}-\bar{R}_e\right)}$$

$$=\frac{1}{\omega_b}\sqrt{\bar{R}_e\left(\frac{K_{\omega\mathbf{B}}}{\mathbf{H}_b}-\bar{R}_e\right)} \text{ henrys per inch} \qquad (25)$$

The two equations will give identical results.

Step 3: Calculate $\bar{L}_0$ from each of the two equations

$$\bar{L}_0=\frac{\bar{R}_e^2+\omega_a^2\bar{L}_e^2}{\omega_a\bar{R}_e\sqrt{(\mathbf{B}_{ma}/\mathbf{B}_a)^2-1}-\omega_a^2\bar{L}_e} \text{ henrys per inch} \qquad (26)$$

$$\bar{L}_0=\frac{\bar{R}_e^2+\omega_b^2\bar{L}_e^2}{\omega_b\bar{R}_e\sqrt{(\mathbf{B}_{mb}/\mathbf{B}_b)^2-1}-\omega_b^2\bar{L}_e} \text{ henrys per inch} \qquad (27)$$

by taking the positive values of the square roots, and then average the values thus obtained.

Step 4: When these calculations have been made for various values of $(d\mathbf{B}/dt)_m$, proceed to plot $\bar{L}_0$, $\bar{L}_e$, and $\bar{R}_e$ as functions of $(d\mathbf{B}/dt)_m$.

References

1. STANDARD TEST METHODS FOR PULSE TRANSFORMER CORES, W. H. Bostick. *Report 722*, Massachusetts Institute of Technology Radiation Laboratory, Cambridge, Mass., May 5, 1945.

2. PULSE GENERATORS (book), G. N. Glasoe, J. V. Lebacqz, editors. McGraw-Hill Book Co., Inc., New York, N. Y., 1948, pp. 631–34.

3. EQUIVALENT CIRCUIT OF A PULSE TRANSFORMER CORE, H. L. Rehkopf. *Report 666*, Massachusetts Institute of Technology Radiation Laboratory, Cambridge, Mass., March 20, 1945.

4. PULSE TRANSFORMER DESIGN AND TEST PROCEDURES, P. R. Gillette, K. W. Henderson, K. Oshima, R. M. Rowe. Final Report on *Contract no. DA 36-039 SC-42645*, pt. III, Stanford Research Institute, Menlo Park, Calif., 1955.

Roll-Welding Precious Metals for Telephone Contacts

A. L. QUINLAN
MEMBER AIEE

THE roll-welding process has been used for years at the Western Electric Company for fabricating tapes for precious metal contacts. A thin strand of palladium a few thousandths of an inch thick and about 0.010 inch wide is roll-welded to a nickel tape of larger section to facilitate handling of the short lengths required in the contact-welding machines. Many millions of contacts are required in the telephone equipment manufactured each year and many contact welders are used.

Several years ago a special roll-welding machine was developed to reduce the contact-welding cost on crossbar switches. With this machine the contact tape is roll-welded to strip stock for springs before they are blanked from the stock. This has proved to be the most economical method of producing precious metal contacts.

More recently the introduction of the Bell System's wire spring relay has required the roll-welding of much heavier precious metal tapes. Here palladium 0.011 inch thick by 0.042 inch wide must be welded to opposite edges of a cupronickel base section of 0.042 by 0.079 inch to make a double-cap tape. In a subsequent welding machine, described in a previous paper,[1] this tape is cut into short lengths as it is welded to the ends of wires to form the common contacts in the wire spring relay. In this construction the problem of the welded portion of the precious metal tape acting as a heavy current-carrying short circuit across the welding electrodes is very pronounced. The large currents involved when high-conductivity metals are welded may soften the metal under the electrodes producing rough-surfaced contacts. A method of calculating the relative importance of the current in the short circuit as compared to the weld current has been developed. This method may be used for evaluating design features such as the use of various metals, weld spacings, and dimensional changes before building a welder.

The Roll Welders

The roll welders used for each of the jobs described herein, welding contact tape to flat stock for blanking spring multiples and welding the heavy sections of contact tape for wire relays, are substantially the same. The details for guiding the stock are varied to suit. The welder line shown in Fig. 1 for the heavy contact tapes is typical. Here the base strand is pulled through the machine by a roll puller located just out of the picture to the right. The speed of the tape through the line is chosen so that the 60-cycle weld energy with pulse peaks 1/120 second apart will be spaced along the tape at the best intervals for adequate strength. The strip is degreased, wire brushed, welded, rolled in a sizing mill, and coiled on a reel. Electronic heat control is provided which sharpens the rise rate of the current pulses and holds constant the power applied to the weld zone. Voltage-drop brushes on the electrodes and a current transformer in the weld circuit furnish a signal to the control unit.

Paper 55-762, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted June 6, 1955; made available for printing August 16, 1955.

A. L. QUINLAN is with the Western Electric Company, Inc., Chicago, Ill.

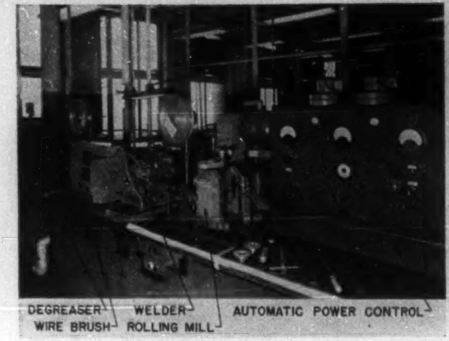


Fig. 1. Roll welder for double-capped tape

Fig. 3. Roll-welding head for spring multiples

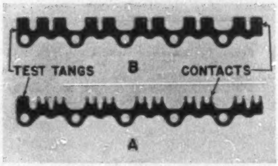


Fig. 2(A). Roll-welded spring multiple. (B). Bar-welded multiple

This unit holds the power at any setting up to 1.5 kw regardless of variations in circuit resistance or line voltage. The cooling water liberally supplied to the weld zone also reduces oxidation.

Ball bearings are used to assure free rotation of the electrodes. The head casting is made of aluminum to reduce its mass. It is free to respond in a vertical direction, guided in ball-bearing ways.

The electrode pressure is set at about 40 pounds. Accurate alignment between the base metal and the precious metal tapes carried in electrode grooves is provided for by a micrometer axial adjustment of the upper electrode head.

Roll-Welding Contacts for Crossbar Switches

The common multiple spring used in crossbar switches is a spring carrying 20 contact points, as shown in Fig. 2. Previous to roll-welding, the springs were made with broad tangs carrying two contact points; see Fig. 2(B). To nest two springs for the economical use of precious metal by the roll-welding method, the tangs were made narrow; see Fig. 2(A). The welding head is shown in Fig. 3. Several 200-foot coils of strip stock are connected by butt-brazing into 1,200- to 1,500-foot coils before roll-welding to avoid many stops of the welding line for rethreading. The take-up reels for the welded coils are made large enough in diameter to prevent the coiled stock from taking a permanent set which would result in bowed parts. The welding line, which has been shown diagrammatically in Fig. 4, operates at about 1,200 feet per hour.

The blanking tool, Fig. 5, is of the multistage type. The first operation is the shearing off of a short length of the tape from the portion of the stock used for the large test tang which must not have a contact; in the second step the precious metal between the tangs which is not required is removed; in the third step the blanking is completed freeing the two spring multiples with all contacts in place.

Since all length dimensions including contact lengths are determined by the die and can be held much more accurately than when welded by the previous individual contact welders, a 20-per-cent reduction in contact length with resulting precious metal saving was realized. Also, the effectiveness of cleaning operations in the line made it possible to use a less expensive brass strip stock instead of nickel silver. Previous attempts to make this change prior to roll-welding have been unsuccessful because of the difficulty of keeping the brass bright until the time of welding.

Weld strength of the contacts is determined with a shear gauge calibrated in pounds. A typical distribution is shown in Fig. 6. For control chart purposes the weekly average strength minus 4 sigma is being held down to a minimum of 20 pounds.

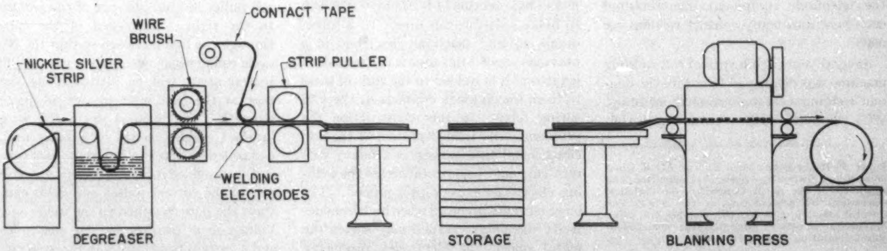


Fig. 4. Diagrammatic sketch of roll-welding line

Table I. Calculated Weld and Shunt Resistances and Currents

	Tape Material					Resistances, Ohms												
	Cap		Base															
Example	Thickness, Inch	Metal	Thickness, Inch	Metal	Weld Space, Inch	(r_1) x10⁻³	(r_2) x10⁻³	(r_3) x10⁻³	(r_4) x10⁻³	$(r_1)^2$ x10⁻⁶	$4(r_1)$ x10⁻³	Weld Branch, Ohms x10⁻³	Resistance of Shunt Branch, Ohms x10⁻³	Total Resistance, Ohms x10⁻³	Shunt Current, Per Cent	Estimated Weld Current, Amperes	Total Current, Amperes	Shunt Current, Amperes
A	0.011	palladium	0.79	cupro-nickel	0.045	2.22	0.725	0.61	0.30	4.93	8.88	1.21	2.795	0.845	30.3	812	1,165	353
B_1	0.011	silver	0.79	cupro-nickel	0.031	0.165	0.725	0.61	0.30	0.027	0.660	1.21	0.437	0.320	73.2	1,200	4,480	3,280
B_2	0.011	silver	0.79	cupro-nickel	0.045	0.248	0.725	0.61	0.30	0.062	0.992	1.21	0.566	0.387	68.5	1,200	3,810	2,610
B_3	0.011	silver	0.79	cupro-nickel	0.062	0.330	0.725	0.61	0.30	1.09	1.320	1.21	0.880	0.510	57.8	1,200	2,850	1,650
B_4	0.005	silver	0.89	nickel	0.060	0.644	0.920	0.17	0.30	0.415	2.58	0.77	1.157	0.462	40.2	1,200	2,010	810
B_5	0.005	silver	0.89	cupro-nickel	0.060	0.644	0.967	0.81	0.30	0.415	2.58	1.41	1.174	0.639	54.7	1,200	2,650	1,450
C_1	0.002 0.009	silver palladium	0.79	cupro-nickel	0.045	0.535	0.725	0.61	0.30	0.286	2.14	1.21	0.944	0.531	53.4	812	1,730	918
C_2	0.002 0.009	silver palladium	0.79	nickel	0.045	0.535	0.697	0.202	0.30	0.286	2.14	0.802	0.933	0.433	46.1	812	1,510	698

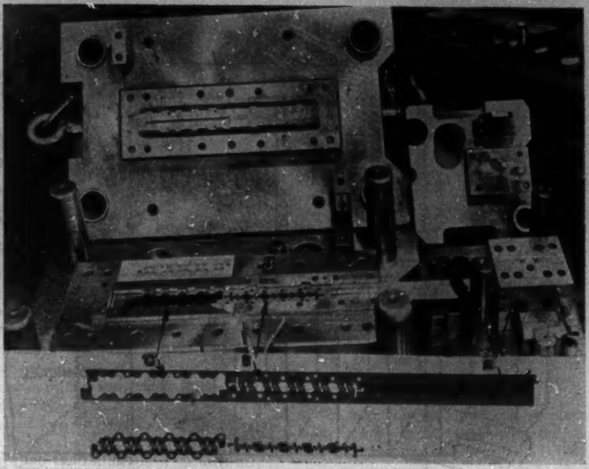


Fig. 5. Blanking die for roll-welded multiples

A—First step: shearing tape from test lug

B—Second step: removing extra contact tape

C—Third step: completing the blanking of two multiples

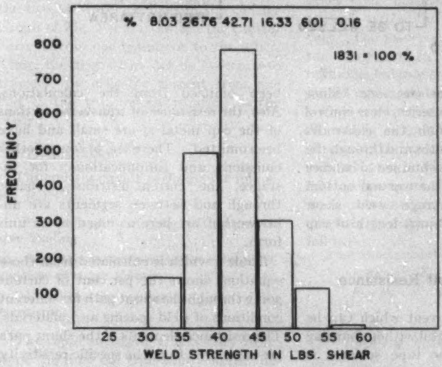


Fig. 6 (left). Weld strength of roll-welded multiple contacts

A Double-Cap Tape

The general construction and dimensions of the double-cap tape are shown in Fig. 7. The cap metal is normally palladium and the base metal cupro-nickel (70-30 per cent). The 0.101-inch dimension between the cap surfaces must be held to close limits as it enters into the limits between "make" and "break" contacts of the finished relay. Also, close limits are necessary to permit ready handling in the automatic percussion welder which cuts contact lengths from this tape and welds them to the ends of the wires in the wire spring relay.[1] The precious metal tapes for the top and bottom sections are formed with weld beads on the surface to be welded. Weld beads concentrate the weld heat in a reduced section, thus facilitating welding by requiring less power. Fig. 7 shows the tape divided into sections or blocks corresponding in length to the space traversed by the welder during each half-cycle of the 60-cycle power used for welding. These block elements are used in the shunt-resistance calculations discussed later.

The welder for this double-capped tape is equipped to feed cap metal simultaneously into grooves in both top and bottom electrodes. The base metal strip is guided in a slot to enter the space between the electrodes. This guide as well as the bottom electrode is electrically insu-

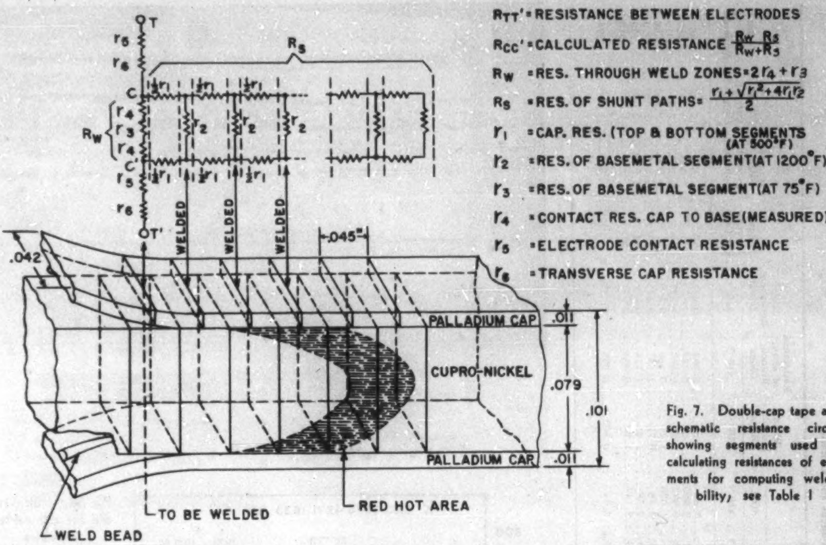

$R_{TT'}$ = RESISTANCE BETWEEN ELECTRODES
$R_{CC'}$ = CALCULATED RESISTANCE $\frac{R_W R_S}{R_W + R_S}$
R_W = RES. THROUGH WELD ZONES = $2r_4 + r_3$
R_S = RES. OF SHUNT PATHS = $\frac{r_1 + \sqrt{r_1^2 + 4r_1r_2}}{2}$
r_1 = CAP. RES. (TOP & BOTTOM SEGMENTS (AT 500°F)
r_2 = RES. OF BASEMETAL SEGMENT (AT 1200°F)
r_3 = RES. OF BASEMETAL SEGMENT (AT 75°F)
r_4 = CONTACT RES. CAP TO BASE (MEASURED)
r_5 = ELECTRODE CONTACT RESISTANCE
r_6 = TRANSVERSE CAP RESISTANCE
T
r_5
r_6
C
R_S
$\frac{1}{2}r_1$
r_2
R_W
r_4
r_3
C'
T'
WELDED
WELDED
WELDED
.045"
.042
PALLADIUM CAP
CUPRO-NICKEL
PALLADIUM CAP
.011
.079
.101
.011
TO BE WELDED
RED HOT AREA
WELD BEAD

Fig. 7. Double-cap tape and schematic resistance circuit showing segments used in calculating resistances of elements for computing weldability, see Table I

Vol.
Pt. 1
1955

lated from the upper electrode. Since the welds are made in series, close control of the heat loss from the electrodes through conduction paths and through the cooling water must be obtained to balance the weld strength in the top and bottom welds. Typical average weld shear strength for an 0.083-inch length of cap weld is 50 pounds.

Calculation of Shunt Resistance

The amount of current which can be conducted into the metal without burning the electrodes or the tape surfaces is limited and the high-conductivity metals discussed herein require currents close to this limit. It is desirable, therefore, to be able to determine in advance what tape designs or modifications can be welded. A method has been worked out of computing the resistance of the direct path between the electrodes as compared to the shunt path, the one through the "just welded" portion of the tape. The high conductivity of these metals makes the shunt-current path very important as will be shown. Note the area trailing the weld zone, Fig. 7, which is red hot because of successive pulses of shunt current. Fig. 7 shows the tape, the schematic circuits, and the equations for the branch resistance. Since the electrode contact resistances r_5 are common to both weld and shunt paths, they have been omitted from the calculations. Also, the resistance of transverse sections of the cap metal r_6 are small and have been omitted. There are, of course, other emissions and simplifications; for instance, the current-distribution paths through and between segments are unknown but are here assumed to be uniform.

Table I which is calculated from those equations shows the per cent of current going through the shunt path for different conditions of weld spacing and materials. The resistance elements of the shunt path are calculated from the specific resistivity of the weld metals and the dimensions of the circuit elements. For instance, example *A*, Table I shows the resistance r_1 calculated for two pieces of palladium each, 0.011 by 0.042 inch in cross section and one weld spacing 0.045 inch long, one piece at the top and one at the bottom at a temperature of 500 degrees Fahrenheit. The resistance r_2, which completes the first shunt path, is calculated for the base metal cross section, 0.042 by 0.045 inch and 0.079 inch long, at 1,200 degrees Fahrenheit which is its estimated operating temperature.

As indicated in the schematic circuit of Fig. 7, the summation of all these shunt paths, each of which represents a weld stitch, is symbolized by R_s. Example *A* shows a power loss of 30.6 per cent through the shunt path. Judged by experience, welds can be made satisfactorily under these conditions. The conditions represented by example C_1 have been welded but are considered to represent limiting conditions with a power loss of 53.4 per cent. A change to the use of nickel base material, as shown by example C_2, would improve weldability as the loss is only 46.1 per cent. Changes to improve the weldability of a silver-capped tape, where the loss is brought down from unweldable 73.2 per cent to 40.2 per cent, are shown as B_1 to B_4 in the table. In this way an advance picture of weldability can be obtained.

Conclusions

The roll-welding and blanking method of making multiple contact springs has resulted in considerable economy and the production of contacts at the lowest cost. Lower-cost base materials can be used with roll-welding because of the thorough cleaning associated with the welding.

Roll-welding, although difficult on high-conductivity metals, can be used to advantage on many applications. A calculation of the effect of the shunt path is desirable to predetermine the weldability of a design.

Reference

1. Automatic Percussion Welding, A. L. Quinlan. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 561-65.

A Symmetrical Transistor Oscillator with Low Second-Harmonic Distortion

W. M. GRIM, JR.
ASSOCIATE MEMBER AIEE

THE oscillator described in this paper was designed to supply the excitation or carrier signal for a second-harmonic type of magnetic modulator. A brief discussion of the modulator will show the necessity for the unusual characteristics of the oscillator; namely, low second-harmonic distortion combined with amplitude and frequency stability.

The magnetic modulator is a device which modulates an alternating carrier with a low-power d-c or direct-voltage input signal. The output signal is twice the frequency of the carrier, proportional to the magnitude of the input and reverses in phase with changes in polarity of the input signal. The minimum detectable power is of the order of 10^{-14} watt, substantially smaller than that of the usual magnetic amplifier.

ol. 74
t. 1
955

To obtain satisfactory operation, the carrier must meet certain critical requirements. Because the modulated output depends on the carrier voltage, the carrier should be of constant amplitude. Perfect carrier balance is not possible, as the modulator output always contains some of the carrier signal; therefore, the carrier should contain as little second harmonic as possible. In the usual modulation system, the desired second-harmonic output is separated from residual signals by means of sharply tuned filter circuits; thus, the carrier should have good frequency stability.

The device described herein has been designed to have these necessary characteristics. Because the power required is small, high power output and high efficiency are not necessary for this application.

Circuit Description

PRINCIPLE OF OPERATION

A conventional push-pull, grounded-base or grounded-emitter oscillator was found to vary greatly in amplitude with changes in room temperature. Thus, an oscillator independent of transistor characteristics was necessary. The device shown in Fig. 1 was adapted from a complex vacuum-tube supply in which a constant current was switched alternately between the ends of a center-tapped tuned transformer by means of vacuum tubes.[1]

The alternating voltage existing across the two 145-turn windings in the emitter circuit of Fig. 1 acts to switch the emitter current from one transistor to the other. Thus, the transistors act as switches to transfer the collector current from one end of the transformer winding to the other.

EMITTER CIRCUIT

The emitter circuit performs the following functions:

1. Permits rapid switching without excessive loading.
2. Supplies constant emitter current regardless of drive voltage.
3. Equalizes the conduction times of the transistors.

The rapid switching is necessary to reduce the time the transistor is in the "active" range of operation, i.e., between cutoff and saturation. During this active range, parameter variations can affect the output. For rapid switching, a large switching voltage is required to insure full emitter-current flow soon after the zero crossing of the switching voltage occurs.

If the center tap of the emitter winding were grounded, the required driving voltage would cause excessive peak currents to flow. These currents would either destroy the transistor or prevent the occurrence of oscillation because of the excessive power drain.

If the center tap is connected to a positive current source, switching is completed when the voltage across the two driving windings is equal to that required by the transistor to conduct the emitter current, about 0.5 volt. The entire current supplied by the source is now supplied to the conducting emitter until the voltage again falls to 0.5 volt. Further increase in the switching voltage produces negligible increase in the emitter current. Thus, the emitter circuit permits rapid switching to be obtained with a minimum of driving power.

The emitter current is determined by the current source as long as the voltage is sufficient to cause switching. Changes in the voltage will, however, cause the transfer time to change. Thus, a change in the developed voltage will not cause a reduction of the driving current and a corresponding further reduction in the output voltage. The magnification of the effects of parameter variations on output usually exhibited in positive feedback systems is therefore reduced greatly.

To have a negligible second-harmonic output, it is necessary that the conduction times of the transistors be equal. If they are of different lengths, a second- (and other even) harmonic component is impressed across the tuned collector circuit. The insertion of capacitor C_2 and resistors R_2 and R_2', as well as the second-harmonic balance control, provides the required equal conduction times.

One explanation of the effects of C_2, R_2, and R_2' is as follows: The equivalent circuit of the basic emitter-switching circuit is shown in Fig. 2(A). The transistors are represented by ideal diodes and an equivalent resistance. The actual currents flowing in the emitters are as follows:

$$i_{e_1}=I_T\frac{R_2}{R_1+R_2}-\frac{E_{sp}}{R_1+R_2}\sin\omega t \tag{1}$$

$$i_{e_2}=I_T\frac{R_1}{R_1+R_2}+\frac{E_{sp}}{R_1+R_2}\sin\omega t \tag{2}$$

or

$$i_{e_1}=I_1-i_s\sin\omega t \tag{3}$$

$$i_{e_2}=I_2+i_s\sin\omega t \tag{4}$$

where

i_{e_1} and i_{e_2} = actual emitter currents, subject to the condition that they can never be negative or greater than I_T
I_T = the total current supplied by the current source
R_1 and R_2 = equivalent transistor input resistances
E_{sp} = peak switching voltage
ω = natural frequency of oscillation

A time plot of this relation is shown in Fig. 2(B) for a case where R_2 is greater than R_1. These resistances and the unequal conducting times can be made equal by adding to the resistance of the smaller one or by use of the more com-

---

Paper 55-740, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted April 2, 1955; made available for printing July 26, 1955.

W. M. GRIM, JR., was formerly with the Electronic Nuclear Instrumentation Group, Servomechanisms Laboratory, Department of Electrical Engineering at the Massachusetts Institute of Technology, and is now with the General Electronic Laboratories, Inc., Cambridge, Mass.

This work was made possible through support extended the Massachusetts Institute of Technology by the Office of Naval Research and the Navy Bureau of Ships.

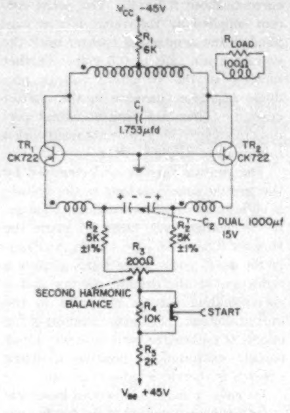


Fig. 1. Symmetrical transistor oscillator. All resistors ±10 per cent, 1 watt unless otherwise specified. Transformer data: core, Arnold Engineering Company's no. A-585079-2, all sections no. 26 wire, collector winding, total inductance, 92 millihenrys, Q, 30 at 400 cycles per second, 500 turns per section, emitter winding, 145 turns per section, and output winding, 100 turns

plicated emitter circuit, shown in Fig. 1, where the d-c circuits are separated by a capacitor and separate identical current sources used to make the currents equal. The second method is superior to the first because it requires no additional driving voltage and the effects of subsequent changes in resistance are minimized.

An alternative explanation of the effect of the two current sources is that, when the circuit is first turned on, a net current flows through the capacitor in the proper direction to charge it to aid the switching voltage in turning on the transistor which conducts the shorter time. This alternative explanation can be used to show that this system tends to provide equal conduction periods even if the switching wave form is distorted so that the zero crossings are not equally spaced.

Unfortunately, the oscillator is not reliably self-starting, chiefly because small oscillations attempting to build up are loaded by the emitter resistances. When oscillations have started, the emitter circuit is loaded by the controlled emitter current, the equivalent of a much larger resistance. Consequently, a normally closed switch is used to decrease momentarily the emitter current and increase the resistance to permit build-up.

COLLECTOR CIRCUIT

The collector supply circuit is designed to provide constant current. In the experimental unit shown, the current is supplied by a battery and resistor combination. The resistance value should be considerably higher than the effective saturated collector resistance, about 50 ohms, to prevent current changes with changes in collector characteristics. Unfortunately, this resistance loads the tank circuit. An increase in this resistance results in a lowering of efficiency because greater supply voltages are required. An inductor in the circuit may be used to reduce the loading without decreasing the efficiency.

Analysis

For analysis, the device may be treated as a tuned center-tapped transformer excited at one end by a square wave of amplitude I_c and loaded by the emitter current wave, the collector circuit resistance, and the load resistor. A possible a-c equivalent circuit is shown in Fig. 3. In the analysis only the components of current and voltage at the resonant frequency are considered.

Analysis of the circuit shows that

$$I_{eff}=\frac{2\sqrt{2}}{\pi}\left(I_c-\frac{I_e}{a}\right) \tag{5}$$

where

I_{eff} = net rms driving current at fundamental frequency
I_c = peak value of collector current
I_e = peak value of emitter current
a = ratio of collector to emitter winding turns

$$E=2Z_{1/2}I_{eff} \tag{6A}$$

$$=2\frac{Z_T}{4}I_{eff} \tag{6B}$$

where

E = total rms voltage of collector winding
Z_T = total impedance between the ends of the tuned circuit, equal to four times the impedance between one end and ground, $Z_{1/2}$.

Combining equations 5 and 6(B) gives

$$E=\frac{\sqrt{2}}{\pi}Z_T\left(I_c-\frac{I_e}{a}\right) \tag{7}$$

For the circuit shown in Fig. 1

$$E=\frac{\sqrt{2}}{\pi}(5{,}340)\left(7.5-\frac{145}{500}10\right)\times 10^{-3}$$
$$=11.0 \text{ volts}$$

This result is about 15 per cent higher than the experimental value. A more complete analysis accounting for the actual switching time might reduce the error at the expense of greater mathematical complexity.

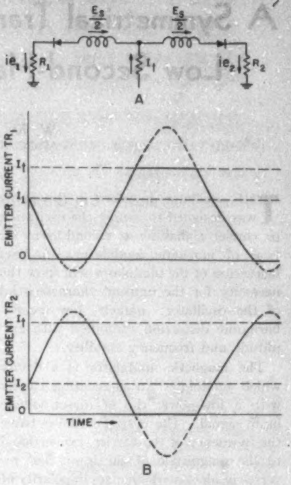


Fig. 2. Emitter circuit

A—Equivalent circuit
B—Calculated emitter current wave forms

Synthesis

To utilize the transistors to full advantage, the equation developed for analysis, when rearranged, can be used to calculate the effective impedance required to develop across the tuned circuit the maximum peak collector voltage for the maximum allowable emitter current. This is

$$Z_T=\frac{\pi}{2}\frac{E_{cp}}{I_{ep}}\left(\frac{a}{a\gamma-1}\right) \tag{8}$$

where

γ = the ratio of collector supply current to emitter supply current to assure collector saturation
E_{cp} = peak collector voltage
I_{ep} = peak emitter current

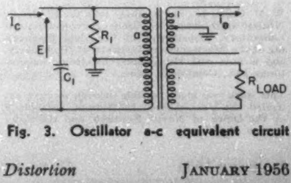


Fig. 3. Oscillator a-c equivalent circuit

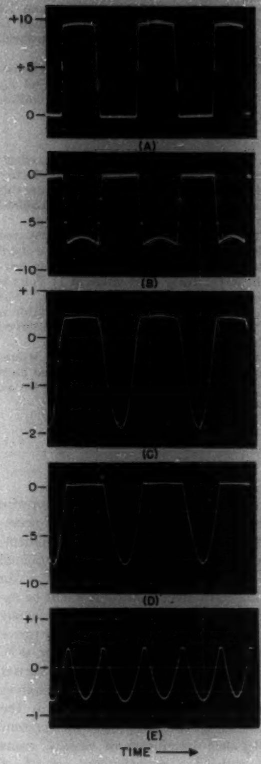


Fig. 4. Oscillator current and voltage wave forms

A—Emitter current, milliamperes
B—Collector current, milliamperes
C—Emitter voltage, volts
D—Collector voltage, volts
E—Reference voltage, volts

For rapid switching, the emitter driving voltage should be about 3 volts rms. For a peak collector voltage of 22 volts, the rms voltage on the tuned circuit will be 15.6 volts. Therefore, the ratio a should be about 5. The resulting driving voltage should be sufficient to cause the switching time to be about 5 per cent of 1 cycle.

The value of γ must be sufficiently less than the grounded-base current gain to insure saturation in spite of possible deterioration of the transistor. A value of 0.75 has been found to be satisfactory.

Thus, for a typical system

$$Z_T=\frac{\pi}{2}\,\frac{22}{10\times10^{-3}}\,\frac{5}{5(0.75)-1}$$
$$=6{,}270 \text{ ohms}$$

This total must be apportioned among the actual reflected load, the tuned circuit impedance, and the collector supply circuit. A possible distribution would be to make the reflected load twice the desired total impedance and to make the tuned circuit impedance and equivalent supply impedance each equal to four times this value. Thus, the reflected load should be 12,550 ohms, the tuned circuit 25,100 ohms, and the actual collector supply 6,270 ohms.

For such a system, the power output would be

$$P=\frac{V^2}{R}=\frac{(15.5)^2}{12{,}550}=19.2\times10^{-3} \text{ watt}$$

The 5,000-ohm resistors in the emitter circuit have proved satisfactory. The capacitor C_2 should be large enough so that only a negligible portion of the emitter driving voltage appears across it.

The total power input from a 50-volt collector supply and a 25-volt emitter supply would be 0.63 watt. The efficiency is 3.1 per cent, substantially higher than that for the experimental model shown in Fig. 1.

Transformer Design

If low-harmonic output is desired, powdered permalloy cores should be used. Arnold Engineering Company's *no. A58-5079-2* has been found satisfactory. With this core, a Q of 30 at 400 cycles can be obtained. If greater Q is desired, a core with larger cross section is required. The number of turns and the wire size can be calculated from the data given in Fig. 1 through the use of the relations that inductance increases as the square of the number of turns, and that the wire area must decrease as the number of turns increases. The transformer used in the experimental model was wound to insure close coupling between all sections of the various windings.

Experimental Results

The actual current and voltage wave forms for the emitter and collector are shown in Fig. 4(A) through (D). The slight dip in the collector current results from the sinusoidal voltage existing between the center tap and the grounded (conducting collector) end of the transformer. The currents do not switch instantaneously, but take approximately 7 per cent of 1 cycle. The switching transient in the emitter and collector currents remains unexplained. Fig. 4(E) shows the voltage existing between one side of the isolating capacitor C_3, Fig. 1, and ground. This voltage could be employed to gate a phase-sensitive detector used to determine the polarity of an input signal to the modulator.

The power output for this nonideal experimental model is 4.4 milliwatts for an input of 0.83 watt. The efficiency is thus 0.53 per cent. The individual emitter and collector dissipations are 2.5 and 1.7 milliwatts respectively. The balance of the power is wasted in the driving circuit and the constant-current determining circuits.

The second-harmonic distortion is, after adjustment, 0.003 per cent, increasing to 0.006 per cent for a 20-degree-centigrade change in one transistor, for a variation in load from 56 to 140 ohms, for changes of collector current from 6.5 to 8.1 milliamperes, changes of emitter current for 9.6 to 10.5 milliamperes, or for the addition of 10 ohms in one emitter circuit.

The variation in output voltage for a 20-degree-centigrade change in transistor temperature is less than 1 per cent. Frequency variation is less than 1/2 per cent over periods consisting of several days each.

Possible Circuit Modifications

The greater portion of the input power is dissipated in the current supply circuits. If this current can be supplied by a circuit with high a-c output impedance and low d-c output impedance, higher efficiencies should be possible. If the independence of transistor characteristics and the low distortion are not desired, a series inductor can be used to replace the current-determining resistors. The emitter circuit capacitor can be omitted. Thus, with proper modification, high efficiencies should be possible for the reason that this configuration extends the advantages of class-C operation to half-cycle conduction.

The collector dissipation per transistor is about 1.7 milliwatts, approximately 1/20th of the rated value. If it can be shown that, for this switching operation, the collector current can be increased until the dissipation rather than the current is the limiting factor, then it should be possible to obtain 1/2 watt of alternating output from a pair of small transistors.

Conclusion

Details have been presented for a transistor circuit which the transistors operate as switches instead of in a linear fashion. The output has extremely low second-harmonic content and constant output. The device has low efficiency. Circuit efficiency has been exchanged for the other advantages. Modifications have been suggested which may increase the efficiency and possibly the power output. Further development will depend on continued investigation of the large-signal parameters of transistors.

Reference

1. The Design of a Practical D.C. Amplifier Based on the Second Harmonic Type of Magnetic Modulator, **S. W. Noble, P. J. Baxandall.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 99, pt. 2, no. 70, Aug. 1952, pp. 327–44.

Discussion

R. F. Shea, (General Electric Company, Schenectady, N. Y.): One item that gave me considerable trouble was the dual 1,000-microfarad capacitor C_3 which consists of two polarized sections. I first assumed that a ground connection had been omitted from the common negative and on this assumption I could not check the results in the paper. Apparently this capacitor is in effect a single nonpolarized capacitor.

In equations 1 to 4, the subscripts e_1 and e_2 should be dropped. I would suggest using $(E_{sp})/(2) \sin \omega t$ instead of $(E_s)/2$ to conform to the text. In Fig. 3, i_e should be I_e.

Other than these typographical corrections, my only other criticism is concerning inconsistency. For example, under "Analysis" $E = 11$ volts, while under "Synthesis" it is 15.6 volts. Further, the contrast between the assumed and actual values is quite large. For example, under "Synthesis" we are given 22 volts peak collector voltage, 15.6 volts rms, 3 volts rms emitter driving voltage. The corresponding values from Fig. 4 are 8 volts, 5.6 volts, and 1.4 volts. Further, 5.6 volts is more than 15 per cent lower than the afore-mentioned 11-volt value.

One other minor point concerns consistency in use of symbols. Both I_s and I_{sp} are employed to indicate peak emitter current.

W. M. Grim, Jr.: Capacitor C_3 is, in effect, a single nonpolarized one. A dual electrolytic capacitor is shown in Fig. 1 because this method was used to obtain the required capacitance in the original model.

The discrepancy between the calculated value of 11 volts and the actual value of 5.6 volts [Fig. 4(D)] occurs because of an oversight in the text. The value of 11 volts is calculated for the unloaded oscillator. The calculated value for the loaded oscillator is 10.1 peak (7.25 rms), closer to the observed value of 8 volts peak (5.6) rms.

The other apparent confusion in values may result because data are presented both for an early experimental model and for an ideal model.

Combined Operation of 24-Channel Cable Carrier and 12-Channel Open-Wire Carrier Systems

G. W. SEARLE
MEMBER AIEE

R. A. SCHAEFER
ASSOCIATE MEMBER AIEE

COMBINED operation of two carrier systems working at different frequencies and on different types of facilities, through frequency-translating equipment, has been used for some time. Combined operation of *K-J* systems was introduced in the mid-thirties. This system makes use of the 12-channel twin-cable *K* carrier operating in the frequency range of 12 to 60 kc, and the 12-channel open-wire *J* system operating in the frequency range of 36 to 144 kc. More recently, 0-*N* combined operation over cable and open wire has become available. The 0-*N* system is a 20-channel single-cable carrier translated into 4-channel open-wire groups in the 0-frequency range. The frequency range of the 20-channel 0-*N* in the cable portion is 40 to 264 kc. That of the open-wire portion extends up to 156 kc. depending upon the frequencies selected for the open-wire section. The 0 system does not co-ordinate in level and frequency with *J*.

With the increasing use of single-cable carrier, the need became apparent in Wisconsin for a system suitable for combined operation in the *N-J* range. Until the present time, equipment designed for operating the cable portion in the *N*-frequency range with frequency translation into the *J* range has not been available.

The alternative of installing terminals back to back at the junction between two carrier systems not designed for combined operation may, of course, always be resorted to, but very substantial economies may be secured where the back to back terminals can be omitted, since that interconnecting equipment amounts to only a fraction of the cost of the two terminals. The Lenkurt *45B* 24-channel single-cable carrier system and the companion *45A* 12-channel open-wire carrier system, together with the necessary frequency-translating equipment, now supplies the need for combined operation in the *N-J* range. Certain problems in level conditions might be expected in translation to the *J*-frequency range of systems which make use of transmitted carrier such as the *N* or 0-*N*. Since *45A* and *B* are suppressed carrier systems, however, they lend themselves readily to combined operation in the *N-J* range. It is not proposed to present a detailed description of these systems but rather to describe the results of combined operation of the two systems working through frequency-translating equipment. The 12-channel open-wire carrier system is described in reference 1.

Description

A Lenkurt *45A* system was installed between Madison and Stevens Point, Wis., in October of 1953. A *45B* system was installed between Milwaukee and Madison, Wis., in May of 1954, at which time the two systems were interconnected for combined operation. This interconnection provides 12 channels from Milwaukee to Madison, and 12 channels from Milwaukee to Stevens Point. This is the arrangement used during the test period. The entire 24 channels will be interconnected when the system goes into regular service.

Fig. 1 shows the route of the cable and open-wire lines as well as the locations of the repeaters. It might be noted that the

---

Paper 55-683, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted March 24, 1955; made available for printing July 19, 1955.

G. W. Searle (deceased) was with the Wisconsin Telephone Company, Milwaukee, Wis., and R. A. Schaefer is with the Wisconsin Telephone Company, Milwaukee, Wis.

l, 74
1
55

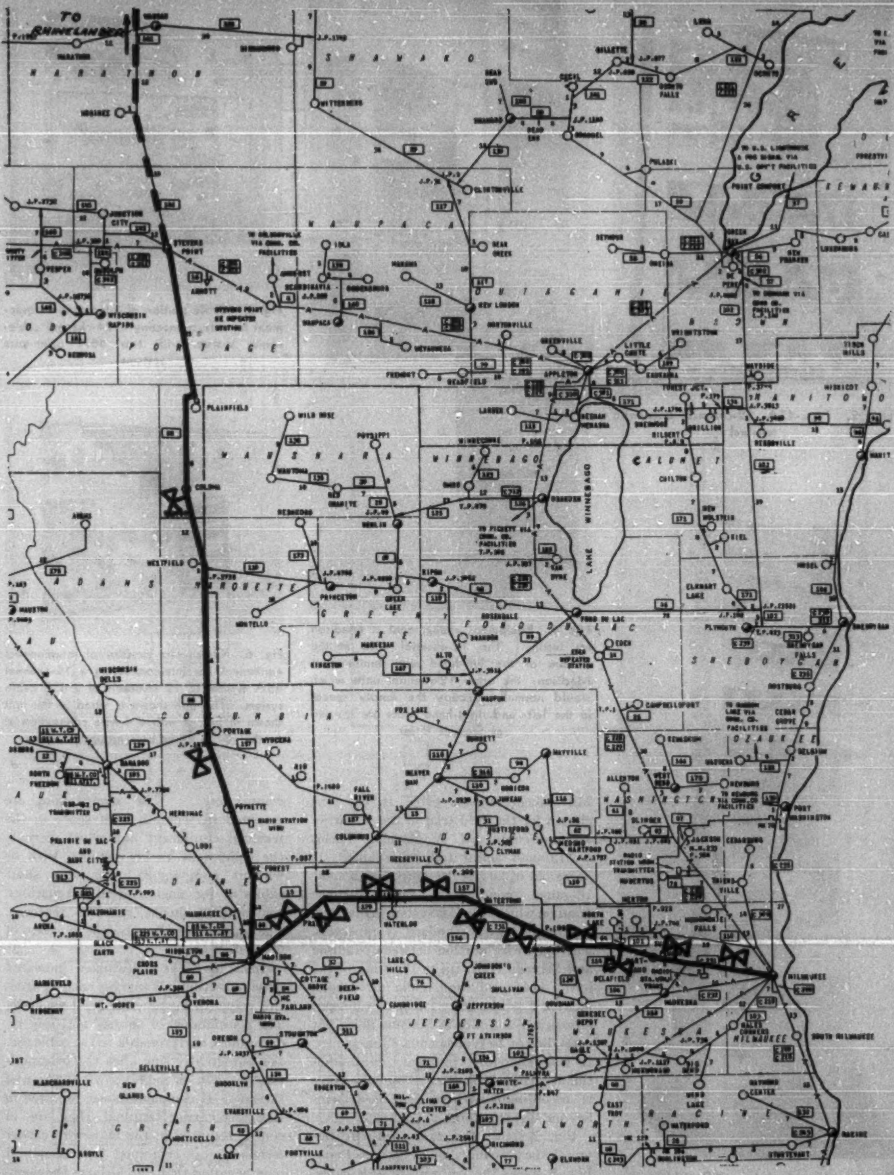

TO RHINELANDER
MARATHON
SHAWANO
STEVENS POINT
WISCONSIN RAPIDS
PORTAGE
WAUPACA
OUTAGAMIE
APPLETON
GREEN BAY
PLAINFIELD
WAUSHARA
WINNEBAGO
LAKE WINNEBAGO
OSHKOSH
CALUMET
COLOMA
WESTFIELD
ADAMS
MARQUETTE
FOND DU LAC
SHEBOYGAN
COLUMBIA
PORTAGE
WISCONSIN DELLS
BARABOO
DODGE
WASHINGTON
OZAUKEE
DANE
MADISON
WATERTOWN
JEFFERSON
WAUKESHA
MILWAUKEE
GREEN
WALWORTH
RACINE

Fig. 1. Cable route from Milwaukee to Madison and open-wire route from Madison to Stevens Point

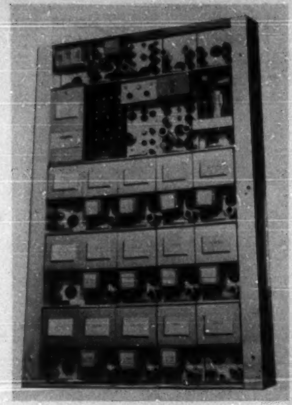


Fig. 2. Type-45A 12-channel open-wire terminal

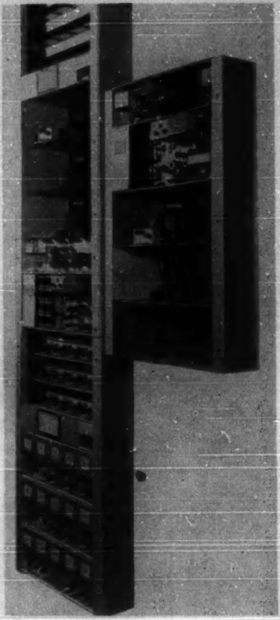


Fig. 4. Equipment arrangement at Madison for testing. The 12-channel units at the bottom of the left-hand rack terminate at Madison; the other 12-channel units which would normally occupy the empty spaces in the left- and right-hand racks are located at Stevens Point

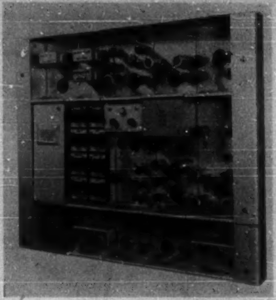


Fig. 5. Cable portion of interconnect equipment for interconnecting a 24-channel cable-carrier system with two 45A open-wire systems

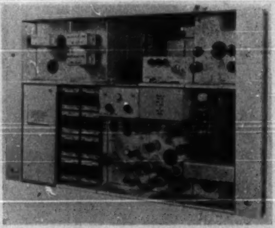


Fig. 6. Open-wire portion of interconnect equipment for interconnecting a 12-channel 45A system to 12 channels of a 45B cable system. The unit shown is wired to the unit shown in Fig. 5 with a 4-wire connection at the 40- to 88-kc base-group frequency

Vol.
Pt. 1
195

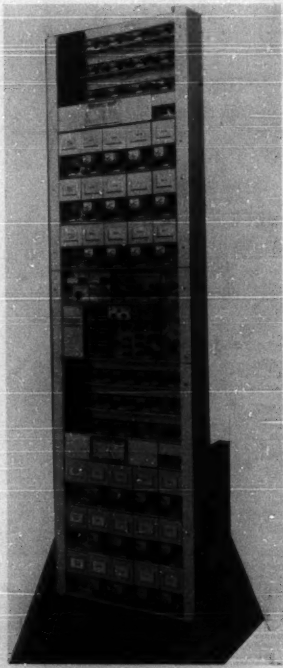


Fig. 3. Milwaukee type-45B cable-carrier terminal. The terminal has 24 channel units plus the necessary common equipment

Milwaukee-Madison cable is the one chosen for the field trial of the *N* carrier and now contains 17 *N* systems. This is a 112-quad aerial toll cable about 25 years old of largely 19 gauges with some 16-gauge conductors. The 19-gauge quad on which the *45B* system is installed is equipped with standard *N* repeaters. The open-wire portion between Madison and Stevens Point up to the time of the initial *45A* installation, had a full complement of *C* assignments. Circuit additions are needed not only to Stevens Point but extending on to Wausau and Rhinelander. It was, therefore, necessary that additional carriers be in the *J*-frequency range to co-ordinate with existing *J* systems beyond Stevens Point.

Fig. 2 shows the *45A* open-wire terminal, the type installed at Stevens Point. Fig. 3 shows the Milwaukee *45B* cable-carrier terminal. All channels are equipped with the Lenkurt miniature compandor. The compandors are mounted in separate shelf space in the terminal equipment bay which permits the equipment to be installed with or without compandors. For very short cables or for single-system installations, compandors should not be required. The compandor advantage is in excess of 22 decibels (db).

Tests made at Milwaukee indicated noise-reduction advantages ranging from this value up to values as great as 35 db. The compandor, of course, not only reduces noise and crosstalk in the cable carrier but greatly simplifies the transposition problem on the open-wire portion of the line. If compandors are required in the cable-carrier terminal, they are, of course, required in the open-wire carrier terminal for combined operation. It may be noted that Fig. 2 shows the open-wire terminal not equipped with compandors, because it was made before com-

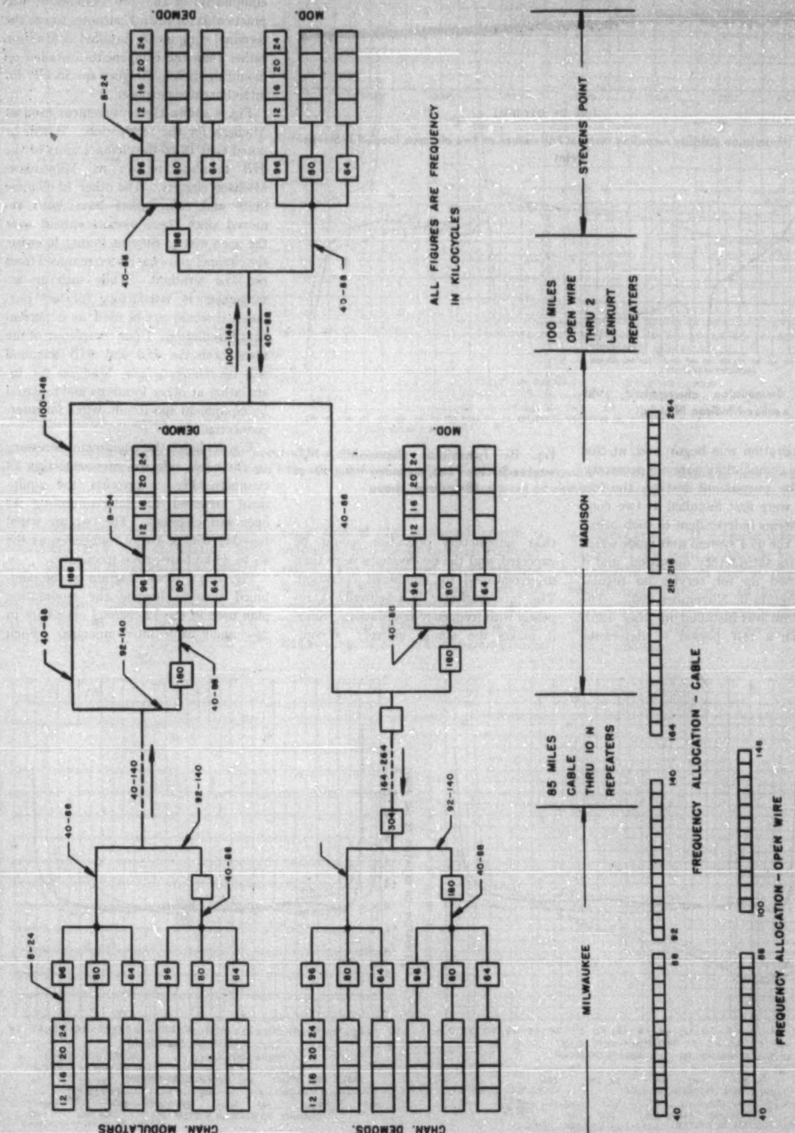

CHAN. MODULATORS
CHAN. DEMODS.
12 16 20 24
96
80
64
8-24
40-88
92-140
180
304
40-140
164-264
100-148
188
DEMOD.
MOD.
ALL FIGURES ARE FREQUENCY IN KILOCYCLES
MILWAUKEE
MADISON
STEVENS POINT
85 MILES CABLE THRU 10 N REPEATERS
100 MILES OPEN WIRE THRU 2 LENKURT REPEATERS
FREQUENCY ALLOCATION - CABLE
40 88 92 140 164 212 216 264
FREQUENCY ALLOCATION - OPEN WIRE
40 88 100 148

Fig. 7. Modulation plan, 12-channel open-wire to 24-channel cable-carrier interconnection

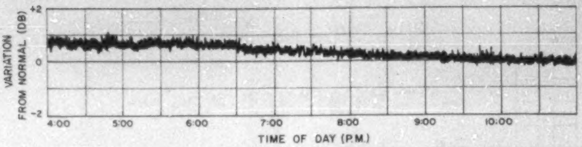


Fig. 8. Transmission stability recording made at Milwaukee on two channels looped at Stevens Point

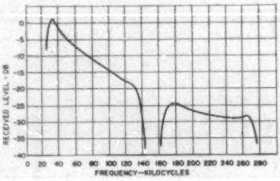


Fig. 9. Transmission characteristic, Milwaukee-Madison N circuit

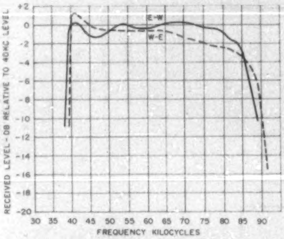


Fig. 10. Transmission characteristic, Milwaukee-Stevens Point, measured at the 40- to 88-kc base group

bined operation was begun and, at that time, the compandors were not necessary.

For the purpose of testing, the two systems were first installed as two complete systems independent of each other. Tests of the *45A* system were made which assured its satisfactory operation, and it was turned up for service on regular traffic circuits in November 1953. The *45B* system was installed in May 1954, was given a test period to determine that satisfactory operation could be expected, and the two systems were then interconnected for combined operation. The interconnector is essentially a repeater with frequency translation. Since it makes use almost entirely of components present in the respective terminal equipments of the two systems, it was practicable for testing purposes to use the terminal equipments installed at Madison rather than the components installed on mounting shelves designed specifically for interconnector purposes.

Fig. 4 shows the arrangement used at Madison for the test period. It will be noted that 12 of the channel units of the *45B* terminal remain as Milwaukee-Madison circuits. The other 12 channel units and compandors have been removed since these circuits extend over the open wire to Stevens Point; likewise, the channel units have been removed from the *45A* terminal. While such an arrangement is satisfactory for test purposes, it would not be used as a permanent installation. Upon completion of the tests, both the *45A* and *45B* terminal will be removed from Madison for installation at other locations and replaced by equipment specifically wired for interconnection.

Fig. 5 shows the equipment necessary on the cable side for interconnecting 24 channels. Fig. 6 presents the equipment required for interconnecting 12 open-wire channels. The two are wired together with a 4-wire connection at the 40 to 88-kc base-group frequency.

Fig. 7 is a block diagram of the combined system showing the modulation plan used in the 12-channel open-wire to 24-channel cable interconnection. Each

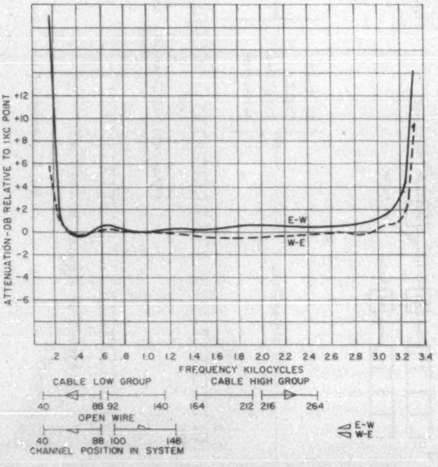


Fig. 11. Frequency response characteristic measured from compressor-in to expandor-out on channel 6, group 1, from Milwaukee to Stevens Point

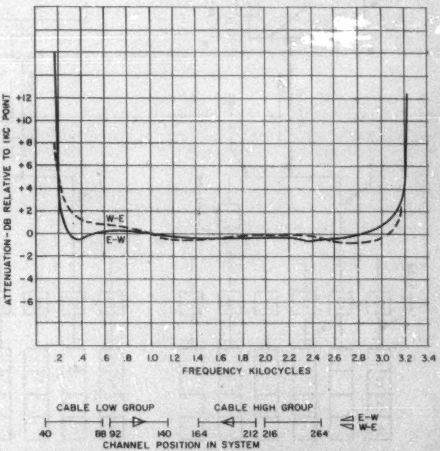


Fig. 12. Frequency response characteristic measured from compressor-in to expandor-out on channel 6, group 2, from Milwaukee to Madison

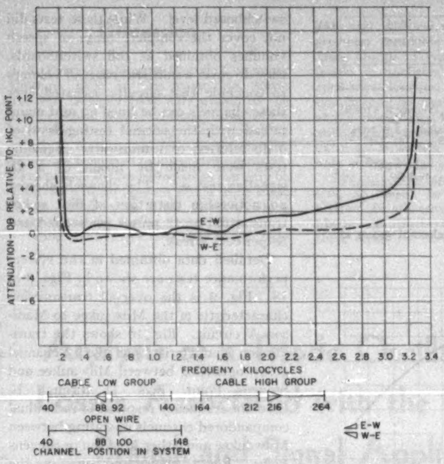


Fig. 13. Frequency response characteristic measured from compressor-in to expandor-out on channel 1, group 1, from Milwaukee to Stevens Point

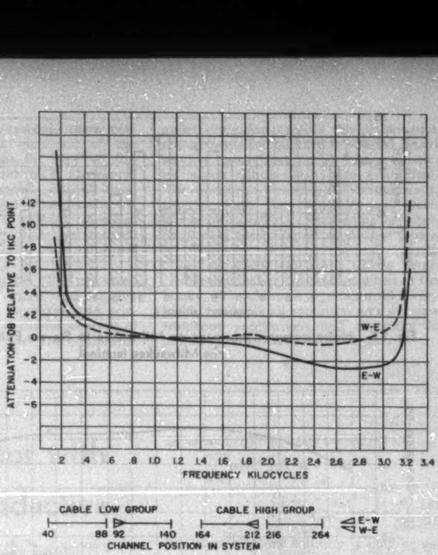


Fig. 15. Frequency response characteristic measured from compressor-in to expandor-out on channel 1, group 2, from Milwaukee to Madison

block in the diagram represents a stage of modulation, with the modulating frequency shown in the block. The desired modulation products are shown on the lines connecting the various modulators.

It will be noted that each 12-channel group on both the open-wire side and the cable side of the combined system is modulated into the 40- to 88-kc frequency range. The 12-channel interconnection is made at this point through pads and amplifiers to provide the proper levels. It will also be noted that there is a frequency band of 8 to 24 kc common to each 4-channel group. It is practical to interconnect at this point to provide dropping channels in groups of four at the intermediate point if desired. In this case the equipment arrangements and modulation plan would be different from those previously shown.

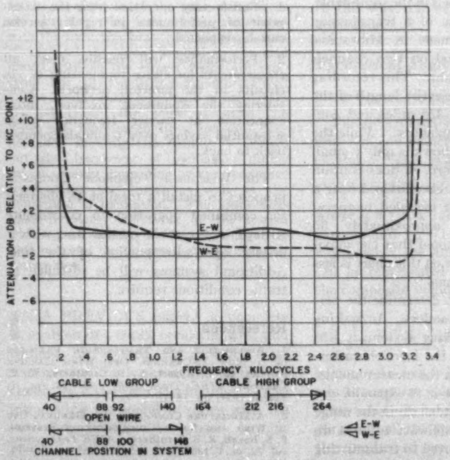


Fig. 14. Frequency response characteristic measured from compressor-in to expandor-out on channel 12, group 1, from Milwaukee to Stevens Point

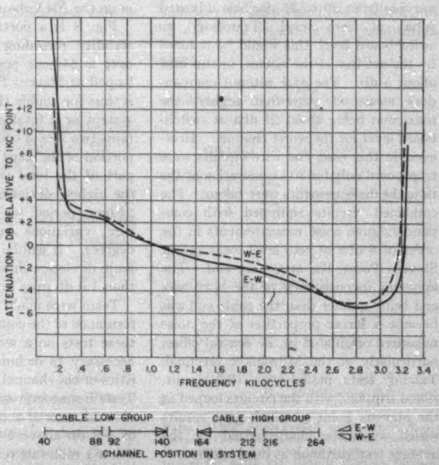


Fig. 16. Frequency response characteristic measured from compressor-in to expandor-out on channel 12, group 2, from Milwaukee to Madison

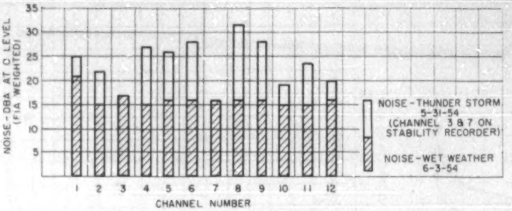


Fig. 17. Noise measurements made on Milwaukee to Stevens Point channel group at the Milwaukee terminal

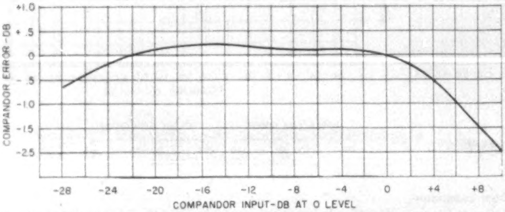


Fig. 18. Compandor error of typical channel, (channel 7, group 2)

Performance

Noise tests were taken on all channels, both on the separate systems and on the systems combined. Noise measurements on both systems were similar. Equipped with compandors, the *45B* gave noise measurements on the various channels varying from 20 to 25 decibels adjusted (dba) at zero level. Translated to switchboard level, this would be reduced by the net loss of the circuit, in this case about 5 db. The *45A* without compandors measured somewhat higher, the maximum being about 21 dba at switchboard level for the worst channel. However, in this case, the wire circuits were transposed only for 30-kc operation at the time the measurements were taken. The combined circuits equipped with compandors gave noise measurements in the order of 8 dba at zero level measured at Stevens Point, the open-wire end. The apparent inconsistency of the open-wire end being quieter than the cable end was because a larger proportion of the noise measured originated in the central office, particularly at the Milwaukee terminal. Talking tests made over the circuit, round trip, i.e., with the circuits looped at the Stevens Point terminal, gave results which would be characterized by the average testboardman as dead quiet.

Crosstalk tests were made on all channels of each system, and on the channels which were interconnected for combined operation. Crosstalk levels were below measurable volumes using the *2B* noise measuring set, and no appreciable crosstalk was apparent from listening tests either on the 24 channels of *45B* or on the *204N* channels in the same cable.

Fig. 8 is a portion of a transmission stability recording made in Milwaukee over a 24-hour period on two channels looped at Stevens Point. This recording is thus for double the entire length of the system or a total of 370 miles and contains two sets of compandors. While the portion of the chart shown is only a small part of the total record, it does contain the highest deviation encountered over a 24-hour period, which included temperature variations of approximately 35 degrees. It will be noted that the maximum deviation shown on this chart is less than 1.0 db from normal.

Tests were made to determine the performance of the compandors. In making these tests on a working system, it was necessary to include the load characteristics of the channel in the measurements. Tests from compressor-in to expandor-out show an error of about 0.3 db in the range of +2 db above 1 milliwatt to −24 db below 1 milliwatt referred to transmitting switchboard level. While these tests did not cover the complete range of speech volumes obtained at toll switchboards, they do cover a sufficient range of volumes to conclude that circuits operated over these channels can be lined up and maintained with the normal testing level of 0 db referred to 1 milliwatt. Signaling tests were made by means of a *2B* signaling test set. The channels showed no noticeable distortion of dial pulses transmitted at 12 pulses per second and 58-per-cent break.

Detailed data obtained in the system performance tests are shown in Figs. 9 to 18. Fig. 9 is the over-all transmission characteristic of the Milwaukee-to-Madison *N* circuit. Fig. 10 shows the transmission characteristics of the 12-channel group operating between Milwaukee and Stevens Point. Figs. 11 through 16 show the frequency response of individual compandored channels operating between Milwaukee and either Madison or Stevens Point. Fig. 17 shows and compares the interfering effects of thunderstorm and wet-weather noise. Fig. 18 shows the error between the compandor input and output for inputs between −28 db and +10 db at zero level.

Conclusion

As a result of rather extensive tests of the *45B* and the *45A* systems separately, as well as tests of the combined operation, it can be concluded that:

1. Satisfactory operation from the standpoint of performance in regular service can be expected.
2. Performance will readily meet all requirements for either terminating or via circuits in the intertoll network. Considering the equipment involved, interconnection at carrier frequency offers substantial savings over terminals operated back to back.

The Wisconsin Telephone Company proposes to install a total of 48 channels for combined operation to co-ordinate with *N* in the cable portion and with *J* in the open-wire portion of the line. Additional systems will be installed as traffic conditions require.

References

1. Application and Transmission Features of a New 12-Channel Open-Wire Carrier System, **K. E. Appert, R. S. Caruthers, W. S. Chaskin.** *AIEE Transactions*, vol. 73, pt. I, March 1954, pp. 18–27.

2. A Miniature Compandor for General Use in Wire and Radio Communication Systems, **F. S. Boxall, R. S. Caruthers.** *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 804–11.

Vol.
Pt. 1
1955

Discussion

J. N. Petrie (Automatic Electric Company, Chicago, Ill.): In this paper, Mr. Searle and Mr. Schaefer have ably presented the results which were achieved (and which have since been repeated many times) by utilizing modern carrier equipment designed for combined or integrated operation. Results such as these, obviously, can only be achieved with equipment which in its basic design incorporates those features which make it suitable for interconnection.

While the authors refer only to the type *45A* and *45B* carrier systems, there now are also available as members of the same family of systems the type *45 BX* (radio channelizing) and *45CB* (4-channel open-wire) carrier systems. All of these systems can be combined or interconnected at the base group (40- to 88-kc) frequencies and at the pregroup (8- to 24-kc) frequencies. (The *45CB* system can be interconnected only at 8 to 24 kc.)

The data presented under "Performance" will, no doubt, be of more than passing interest to discerning engineers and telephone organizations. A careful analysis of these data will reveal that the equipment under discussion exceeds the performance requirements for military full-period leased services, in respect to all characteristics for which criteria have been established. The superior performance of this equipment, together with the economies and flexibilities obtainable by combining or interconnecting (dropping or inserting of groups of channels), its miniaturized plug-in design of units which are largely interchangeable between all four members of the family, all would seem to recommend it very highly for use in all important long distance telephone applications and especially in military full-period leased facilities.

---

ol. 74
t. 1
955

Analysis of a Triode Oscillator with Losses Associated with the Inductive Branch and Signal Applied in Grid Circuit

ABD EL-SAMIE MOSTAFA
NONMEMBER AIEE

Synopsis: By transforming the nonlinear equations of a triode oscillator, with and without an applied signal, to linear ones with periodically varying coefficients, the complete solutions of oscillator equations to the first power of the tube parameters are obtained, the losses being associated with the inductive branch of the oscillatory circuit and the signal applied in the grid circuit. When the circuit losses are relatively low, the resulting harmonics as well as the fundamental amplitude of oscillation are practically the same as when the losses are represented by a constant conductance shunting the oscillatory circuit, but the frequency of oscillation differs by an amount depending on the resistance representing the losses as well as the amplitude of oscillation. The amplitude-phase-frequency relations of the system within and outside the locking range are obtained and the ranges of synchronization determined.

IN A PREVIOUS paper,[1] a complete analysis of a triode oscillator with and without an applied signal was given, representing the losses in the oscillatory circuit by a constant conductance G and introducing the signal in series with the inductive branch. As stated in the paper, G cannot be considered as absolutely constant. The losses are actually associated with the inductive branch, and the signal is usually applied in the grid circuit. The present work gives a complete analysis of a triode oscillator with and without an applied signal, with the losses associated with the inductance and the signal applied in the grid circuit. In this way, the actual practical case is approached.

Theoretical Analysis

FREE OSCILLATIONS

The differential equation for the current i in the inductive branch of the oscillator in Fig. 1 is given by

$$d^2i/dt^2+(r/L)di/dt+\omega_0^2(i_p+i)=0 \qquad (1)$$

where i_p is the a-c component of the plate current and $w_0^2=1/LC$. The resistance r represents the losses in the circuit; i_p can be represented as a power series of the composite voltage e, thus

$$i_p=\sum_{n=1} a_ne^n=e\sum_{n=1}a_ne^{n-1}$$

where $e=v+\mu v_g=L(1-\mu M/L)di/dt+ri$, v is the output voltage and v_g is the a-c component of the grid voltage. The upper limit of the summation depends on the number of terms required in the expansion for satisfactory representation, but generally it is infinity; μ = amplification factor of the tube.

Substitution in equation 1 gives

$$d^2i/dt^2+(1/L)[r+(L/C)(1-\mu M/L)\times \sum_{n=1}a_ne^{n-1}]\,di/dt+\omega_0^2\times \left(1+r\sum_{n=1}a_ne^{n-1}\right)i=0 \qquad (2)$$

For an oscillatory solution, i may be assumed to be of the form

$$i=I_t\cos\int\omega dt+\sum_{m=2}^{\infty}I_{tm}\cos\left(m\int\omega dt+\theta_m\right)$$

where I_t and $\omega/2\pi$, being very slow functions of time compared with the frequency of oscillation, are the fundamental amplitude and frequency at any instant t; I_{tm} and θ_m are also slow functions of time and represent the amplitudes and phases of the harmonics respectively. For the solution of equation 2 to the first power of the tube parameters a_n, it is sufficient to substitute $I_t\cos\int\omega dt$ for i in $\sum_{n=1} a_ne^{n-1}$ of equation 2. Substitution gives

$$d^2i/dt^2+(R_{0t}/L)\left[1+\sum_{m=1}^{\infty}h_m\times \cos\left(m\int\omega dt+\theta_t\right)\right]di/dt+ \omega_{0_t}^2\left[1+\sum_{m=1}^{\infty}k_m\cos m\left(\int\omega dt+\theta_t\right)\right]\times i=0 \qquad (3)$$

where

$$R_{0_t}=r+(L/C)(1-\mu M/L)h_{0_t} \qquad (4)$$

$$\omega_{0_t}^2=\omega_0^2(1+rh_{0_t}) \qquad (5)$$

$$h_{0_t}=\sum_{n=1}\{a_n/2^{n-1}[(n-1)/2]!\}\times [(n-1)(n-2)(n-3)\ldots \text{ up to } (n-1)/2 \text{ factors}]E_t^{n-1} \qquad (6)$$

Paper 55-705, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3-7, 1955. Manuscript submitted May 17, 1955; made available for printing September 22, 1955.

ABD EL-SAMIE MOSTAFA is with Alexandria University, Alexandria, Egypt.

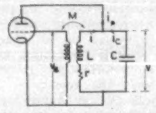


Fig. 1. Tuned plate oscillator, losses being associated with the inductive branch and polarized potentials omitted

(n being odd)

$$R_{0t}h_m=(L/C)(1-\mu M/L)\sum_{\substack{p=(m+1)(m+3),\\(m+5),\ldots}}\times\{a_p/2^{p-2}[(p-m-1)/2]!\}\times[(p-1)(p-2)(p-3)\ldots\text{ up to }(p-m-1)/2\text{ factors}]E_t^{p-1}\quad(7)$$

$$(1+rh_{0t})k_m=R_{0t}h_m[r/(L/C)(1-\mu M/L)]\quad(8)$$

$$E_t=I_t\omega L\sqrt{(1-\mu M/L)^2+1/Q_1^2}\quad(9)$$

$$Q_1=\omega L/r_1,\ r_1=r+L(1-\mu M/L)\times(1/I_t)dI_t/dt,\text{ and }\tan\theta_t=\omega L(1-\mu M/L)/r_1\quad(10)$$

In this way, the nonlinear differential equation 2 has been transformed to a linear one with the coefficients of both di/dt and i periodically varying.

For the solution of equation 3 to the first power of the tube parameters, the methods of superposition and successive approximation are adopted and given in the Appendix. The current i will be given by

$$i=A_s\epsilon^{\int-(R_{0t}/2L)(1+h_2/2)\,dt}\Big\{\cos\int\omega dt-(R_{0t}h_1/2\omega L)(1/\sin\theta_t)+(R_{0t}/2\omega L)(1/\sin\theta_t)\times\sum_{m=2}^{\infty}[1/(m^2-1)](h_{m-1}+h_{m+1})\times\cos m\Big(\int\omega dt+\theta_t\Big)\Big\}\quad(11)$$

$$=I_t\cos\int\omega dt+I_{\text{d-c}t}+\sum_{m=2}^{\infty}I_{mt}\cos m\Big(\int\omega dt+\theta_t\Big)\quad(11\text{A})$$

where A_s is a constant depending on initial conditions and R_{0t} and $R_{0t}h_m$ are to be obtained from equations 4, 6, and 7. The instantaneous angular frequency of oscillation ω is also given, in the Appendix, by

$$(\omega^2-\omega_0^2)/\omega_0^2=[r/(L/C)(\mu M/L-1)]\times[r-R_{0t}(1+h_2/2)]\quad(12)$$

$$=(r/\rho)+r\sum_{p=3}\{pa_p/2^{p-1}\times[(p-1)/2]!\}[(p-1)(p-2)\times(p-3)\ldots\text{ up to }(p-3)/2\text{ factors}]E_t^{p-1}\quad(12\text{A})$$

(p being odd) where $\rho(=1/a_1)$ is the plate resistance of the tube. It is clear from equation 11 that the instantaneous amplitude I_t builds up from the small starting value A_s to the stationary value I according to the relation

$$I_t=A_s\epsilon^{\int-(R_{0t}/2L)(1+h_2/2)\,dt}$$

giving

$$(1/I_t)dI_t/dt=-(R_{0t}/2L)(1+h_2/2)$$

Stationary state is attained when $dI_t/dt=0$, thus $h_2=-2$; substitutions from equations 4, 6, and 7 give

$$\alpha_1-G=(1-\mu M/L)\sum_{p=3,5,\ldots}\{pa_p/2^{p-1}\times[(p-1)/2]!\}[(p-1)(p-2)(p-3)\ldots\text{ up to }(p-3)/2\text{ factors}]E^{p-1}\quad(13)$$

$$=(1-\mu M/L)[(3/4)a_3E^2+(5/8)a_5E^4+(35/64)a_7E^6+(63/128)a_9E^8+(231/512)a_{11}E^{10}+\ldots]\quad(13\text{A})$$

where

$$\alpha_1=(1/\rho)(\mu M/L-1),\ G=Cr/L,\ E=I\omega L\times\sqrt{(1-\mu M/L)^2+1/Q^2},\ Q=\omega L/r$$

and μ is the amplification factor of the tube.

Equation 13 gives the stationary possible fundamental amplitudes E of the composite voltage. If the losses of the circuit are low so that r^2 can be neglected compared with $(1-\mu M/L)^2\omega^2L^2$, equation 13 will be the same as equation 7 of reference 1, only $a_p(1-\mu M/L)^p$ in equation 13 is to be replaced by a_p.

At the stationary state, equations 11 and 12 reduce to

$$i=I\Big\{\cos\omega t-(R_0h_1/2\omega L)(1/\sin\theta)+(R_0/2\omega L)(1/\sin\theta)\sum_{m=2}^{\infty}[1/(m^2-1)]\times(h_{m-1}+h_{m+1})\cos m(\omega t+\theta)\Big\}\quad(11\text{B})$$

$$=I\cos\omega t-I_{\text{d-c}}+\sum_{m=2}I_m\cos m(\omega t+\theta)\quad(11\text{C})$$

and

$$(\omega^2-\omega_0^2)/\omega_0^2=r^2/(L/C)(\mu M/L-1)\quad(12\text{B})$$

$$=(r/\rho)+r\sum_{p=3}\{pa_p/2^{p-1}\times[(p-1)/2]!\}[(p-1)(p-2)(p-3)\ldots\text{ up to }(p-3)/2\text{ factors}]E^{p-1}\quad(12\text{C})$$

where

$$I_m=\text{amplitude of the }m\text{th harmonic}=I(R_0/2\omega L)(1/\sin\theta)[1/(m^2-1)]\times(h_{m-1}+h_{m+1})\quad(14)$$

$$I_{\text{d-c}}=\text{d-c component}=I(R_0h_1/2\omega L)\times(1/\sin\theta),\text{ and }\tan\theta=(\omega L/r)(1-\mu M/L)$$

R_0h_m is the value of $R_{0t}h_m$ obtained from equation 7 after substituting E for E_t.

At critical coupling, $E_t\approx0=dI_t/dt$ and equations 12(A) and 13(A) reduce respectively to

$$(\omega^2-\omega_0^2)/\omega_0^2=r/\rho\text{ and }\mu M/L-1=Cr\rho/L$$

and which are the well-known conditions for sustained oscillations. It is clear that the frequency of oscillation at critical coupling is higher than that given by equation 12(B). The output voltage

$$v=Ldi/dt+ri$$

$$=V\cos(\omega t+\beta)-I_{\text{d-c}}r-\sum_{m=2}mI_m\omega L\times\sqrt{1+(1/mQ)^2}\sin[m(\omega t+\theta)-\beta_m]\quad(15)$$

where

$$V=I\sqrt{r^2+\omega^2L^2},\quad\tan\beta=\omega L/r=Q,\text{ and }\tan\beta_m=1/mQ$$

Delaying the time t by an amount θ/ω, equation 15 becomes

$$v=V\cos(\omega t-\theta+\beta)-I_{\text{d-c}}r-\sum_{m=2}mI_m\omega L\sqrt{1+(1/mQ)^2}\times\sin(m\omega t-\beta_m)\quad(15\text{A})$$

Again, if the circuit losses are low so that r can be neglected compared with $\omega L(\mu M/L-1)$, then $\theta\approx-90$ degrees, $\beta\approx90$ degrees, and $\beta_m\approx0$ degrees; equation 15(A) reduces therefore to

$$v=-V[\cos\omega t-(V_{\text{d-c}}/V)-\sum_{m=2}\times(V_m/V)\sin m\omega t]\quad(15\text{B})$$

where

$$V_m/V=(1-\mu M/L)(1/2\omega C)[m/(m^2-1)]\times\Big((a_m/2^{m-2})E^{m-1}+\sum_{\substack{p=(m+2),(m+4),\\(m+6),\ldots}}\{pa_p/2^{p-2}[(p-m)/2]!\}[(p-1)(p-2)\ldots\text{ up to }(p-m-2)/2\text{ factors}]E^{p-1}\Big)\quad(16)$$

and

$$V_{\text{d-c}}/V\approx(R/2)\sum_{p=2,4,\ldots}\{a_p/2^{p-2}[(p-2)/2]!\}[(p-1)(p-2)\ldots\text{ up to }(p-2)/2\text{ factors}]E^p$$

$$=R[(a_2/2)E^2+(3/8)a_4E^4+(5/16)a_6E^6+(35/128)a_8E^8+\ldots]$$

The harmonics given by equation 16 are the same as those of equation 9(A) of reference 1, only $a_p(1-\mu M/L)^p$ in equation 16 is to be replaced by a_p.

The composite voltage

$$e=L(1-\mu M/L)di/dt+ri$$

$$=E\cos(\omega t+\theta)-I_{\text{d-c}}r-\sum_{m=2}mI_m\omega L\times(1-\mu M/L)\sqrt{1+[1/mQ(1-\mu M/L)]^2}\sin[m(\omega t+\theta)-\gamma_m]$$

where $\tan\gamma_m=1/mQ(1-\mu M/L)$

It is clear that if r is neglected compared with $\omega L(1-\mu M/L)$, then $e\approx(\mu M/L-1)v$. The effects of the harmonics on the frequency of oscillation can be obtained following the same lines as in reference 1. Consider the harmonics, the instantaneous angular frequency of oscillation will be given by

$$(\omega^2-\omega_0^2)\omega_0^2=[r^2/(L/C)(\mu M/L-1)]\times [1+(2Q/\omega)(1/I_t)dI_t/dt]-\sum_{m=2}(m^2-1)(V_{m_t}^2/V_t^2) \qquad (17)$$

V_{m_t}/V_t is the value of V_m/V given by equation 16 after substituting E_t for E.

At stationary state, equation 17 reduces to

$$(\omega^2-\omega_0^2)/\omega_0^2=\{r^2/(L/C)(\mu M/L-1)\}-\sum_{m=2}(m^2-1)(V_m^2/V^2) \qquad (17A)$$

SIGNAL APPLIED

The differential equation for the current i in the inductive branch of the oscillator of Fig. 2 is given by

$$d^2i/dt^2+(r/L)di/dt+\omega_0^2(i_p+i)=0 \qquad (18)$$

The a-c component i_p of the plate current can also be represented as a power series of the composite voltage e, thus

$$i_p=\sum_{n=1}^{\infty}a_ne^n=e\sum_{n=1}a_ne^{n-1}$$

However, in this case the composite voltage e is given by

$$e=L(1-\mu M/L)di/dt+ri+\mu e_s$$

where e_s is an external signal given by
$e_s=E_s\cos\omega_st$
Substitution in the series of i_p gives

$$i_p=\left[L(1-\mu M/L)\sum_{n=1}a_ne^{n-1}\right]di/dt+\left(r\sum_{n=1}a_ne^{n-1}\right)i+\mu E_s\cos\omega_st$$

Let the current i be given by

$$i=I_t\cos\int\omega dt+\sum_{m=2}^{\infty}I_{t_m}\cos\left(m\int\omega dt+\theta_m\right)$$

where I_t, I_{t_m} and θ_m are the fundamental amplitude and the amplitudes and phases of the harmonics respectively; they are slowly varying functions of time compared with the frequency of oscillation. Where ω is the instantaneous angular frequency of oscillation, let it be given by $\omega=\omega_s+d\phi_t/dt$, and $\int\omega dt=\omega_st+\phi_t$. Also ϕ_t and $d\phi_t/dt$ are slowly varying functions of time compared with the frequency of oscillation.

For the solution of equation 18 to the first power of the tube parameters, it is sufficient to substitute $I_t\cos(\omega_st+\phi_t)$ for i in the composite voltage e of the series of i_p. Substitution gives

$$e=[L(1-\mu M/L)-(\mu E_s/\omega I_t)\sin\phi_t]di/dt+[r+(\mu E_s/I_t)\cos\phi_t]i$$

Substitution in equation 18 gives

$$d^2i/dt^2+(1/L)\left[r+\{(L/C)(1-\mu M/L)-(\mu E_s/\omega CI_t)\sin\phi_t\}\sum_{n=1}a_ne^{n-1}\right]di/dt+\omega_0^2\left[1+\{r+(\mu E_s/I_s)\cos\phi_t\}\sum_{n=1}\times a_ne^{n-1}\right]i=0 \qquad (19)$$

If equations 19 and 2 are compared, and if the expressions of e in the two equations are compared, it is clear that equation 19 is similar to equation 2, only $[L(1-\mu M/L)-(\mu E_s/\omega I_t)\sin\phi_t]$ replaces $L(1-\mu M/L)$, and $[r+(\mu E_s/I_t)\cos\phi_t]$ replaces r except in the expression of R_{0_t}. Therefore the current i of equation 19 is

$$i=A_se^{\int-(R_{0t}/2L)(1+h_2/2)dt}\{\cos\int\omega dt-(R_{0_t}h_1/2\omega L)(1/\sin\phi_t)+(R_{0_t}/2\omega L)\times(1/\sin\phi_t)\sum_{m=2}^{\infty}[1/(m^2-1)](h_{m-1}+h_{m+1})\cos m(\textstyle\int\omega dt+\theta_t)\} \qquad (20)$$

where A_s is a constant depending on initial conditions and R_{0_t}, $R_{0_t}h_m$, and θ_t are given by

$$R_{0_t}=r+(L/C)\{(1-\mu M/L)-(\mu E_s/\omega LI_t)\sin\phi_t\}h_{0t}$$

$$h_{0_t}=\sum_{n=1}[a_n/2^{n-1}[(n-1)/2]!]\times[(n-1)(n-2)(n-3)\ldots\text{ up to }(n-1)/2\text{ factors}]E_t^{n-1}$$

$$R_{0_t}h_m=(L/C)\{(1-\mu M/L)-(\mu E_s/\omega LI_t)\sin\phi_t\}\sum_{\substack{p=(m+1),(m+3)\\(m+5),\ldots}}\times[a_p/2^{p-2}[(p-m-1)/2]!]\times\{(p-1)(p-2)(p-3)\ldots\text{ up to }(p-m-1)/2\text{ factors}\}E_t^{p-1}$$

$$E_t\approx I_t\omega L\sqrt{\{(1-\mu M/L)-(\mu E_s/\omega LI_t)\times\sin\phi_t\}^2+[(1/Q_1)+(\mu E_s/\omega LI_t)\cos\phi_t]^2}$$

$$Q_1=\omega L/r_1$$

$$\tan\theta_t=[(1-\mu M/L)-(\mu E_s/\omega LI_t)]/[(1/Q)+(\mu E_s/\omega LI_t)\cos\phi_t]$$

The instantaneous angular frequency of oscillation ω is given by $\omega=\omega_s+d\phi_t/dt$ and

$$d\phi_t/dt=-Z+(\omega_0/2)\{[(1/Q)E_{s_1}\cos\phi_t+(r^2C/L)I_t\omega L/(\mu M/L-1)]/(I_t\omega L+E_{s_1}\sin\phi_t)\}[1+(2Q/\omega I_t)dI_t/dt]-\sum_{m=2}^{\infty}(m^2-1)V_{m_t}^2/V_t^2 \qquad (21)$$

where

$$Z=\omega_s-\omega_0$$

$$E_{s_1}=\mu E_s/(\mu M/L-1)$$

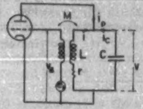


Fig. 2. Tuned plate oscillator, same as in Fig. 1, but with a signal applied in the grid circuit

$$V_{m_t}/V_t=[(1-\mu M/L)-(\mu E_s/\omega LI_t)\sin\phi_t]\times(1/2\omega C)[m/(m^2-1)]\left((a_m/2^{m-2})\times E_t^{m-1}+\sum_{\substack{p=(m+2),(m+4),\\(m+6),\ldots}}\{pa_p/2^{p-2}\times[(p-m)/2]!\}[(p-1)(p-2)\ldots\text{ up to }(p-m-2)/2\text{ factors}]E_t^{p-1}\right)$$

and

$$(1/I_t)(dI_t/dt)=-(R_{0t}/2L)(1+h_2/2)=(1/2C\rho)\{(\mu M/L-1)+(\mu E_s/\omega LI_t)\times\sin\phi_t\}-r/2L+(1/2C)[(\mu M/L-1)+(\mu E_s/\omega LI_t)\sin\phi_t]\sum_{p=3,5,\ldots}\{pa_p/2^{p-1}\times[(p-1)/2]!\}[(p-1)(p-2)\ldots\text{ up to }(p-3)/2\text{ factors}]E_t^{p-1} \qquad (22)$$

Equations 21 and 22 are the amplitude-phase-frequency relations of a triode oscillator with an applied signal in its grid circuit, whether the system is within or outside the synchronizing range. Graphical integration using the step-by-step method is usually adopted for their solutions. If terms up to a_3 only are considered, equation 22 reduces to

$$dI_t/dt=(\epsilon/2L)\{x-(x^3/V_0^2)+mE_{s_1}\sin\phi_t-[x(I_t\omega L)^2/V_0^2][(E_{s_1}/I_t\omega L)^2\cos^2\phi_t+(1/Q^2)+(2/Q)(E_{s_1}/I_t\omega L)\cos\phi_t]\}$$

where

$$\epsilon=(\alpha_1-G)/\omega C$$

$$m=G/(\alpha_1-G)$$

$$x=I_t\omega L+E_{s_1}\sin\phi_t$$

$$V_0^2=(\alpha_1-G)/(-3a_3/4)(\mu M/L-1)^3$$

Stationary synchronized state is attained when

$$dI_t/dt=0=d\phi_t/dt$$

If $dI_t/dt=0=d\phi_t/dt$, and $I_t=I$, $\phi_t=\phi$, $E_t=E$, $V_t=V$, and $V_{m_t}=V_m$ in equations 21 and 22, rearranging

$$2Z/\omega_0=(1/Q)[E_{s_1}\cos\phi/(I\omega L+E_{s_1}\times\sin\phi)]+r\left((1/\rho)+\sum_{p=3,5,\ldots}\{pa_p/2^{p-1}[(p-1)/2]!\}[(p-1)(p-2)\ldots\text{ up to }(p-3)/2\text{ factors}]\right)E^{p-1}-\sum_{m=2}^{\infty}(m^2-1)V_m^2/V^2 \qquad (23)$$

ol. 74
. 1
55

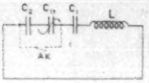


Fig. 3. Connection diagram for equation 27

$$1+m\,\frac{(E_{s_1}/I\omega L)\sin\phi}{1+(E_{s_1}/I\omega L)\sin\phi}=\frac{\displaystyle\sum_{p=3,5,\ldots}\{pa_p/2^{p-1}[(p-1)/2]!\}[(p-1)\times(p-2)\ldots\text{ up to }(p-3)/2\text{ factors}]E^{p-1}}{\displaystyle\sum_{p=3,5,\ldots}\{pa_p/2^{p-1}[(p-1)/2]!\}[(p-1)\times(p-2)\ldots\text{ up to }(p-3)/2\text{ factors}]E_0^{p-1}}\quad(24)$$

where E_0 would be the amplitude of oscillation were the signal absent and it can be obtained from equation 13. If terms up to a_3 only are considered, equations 23 and 24 reduce to

$$2Z/\omega_0=(1/Q)\{(E_{s_1}\cos\phi)/(I\omega L+E_{s_1}\times\sin\phi)\}+[r/(\mu M/L-1)](\alpha_1-G)\times(1+m-E^2/E_0^2)-3V_2^2/V^2-8V_3^2/V^2\quad(23A)$$

$$1+m(E_{s_1}/I\omega L)\sin\phi/[1+(E_{s_1}/I\omega L)\sin\phi]=E^2/E_0^2\quad(24A)$$

Elimination of ϕ between equations 23(A) and 24(A) gives

$$2Z/\omega_0+(3V_2^2/V^2+8V_3^2/V^2)-(\epsilon/Q)\times[1/(\mu M/L-1)](1+m-E^2/E_0^2)+\epsilon^2(1-E^2/E_0^2)^2=[E_{s_1}^2/(I\omega L)^2]\epsilon^2\times(1+m-E^2/E_0^2)^2\quad(25)$$

where

$$E^2=(I\omega L)^2(\mu M/L-1)^2([1+(E_{s_1}/I\omega L)\times\sin\phi]^2+\{[1/Q(\mu M/L-1)]+(E_{s_1}/I\omega L)\cos\phi\}^2)$$

In most practical cases

$$E^2\approx(I\omega L)^2(\mu M/L-1)^2[1+(E_{s_1}/I\omega L)\times\sin\phi]^2\approx V^2(\mu M/L-1)^2[1+(E_{s_1}/V)\sin\phi]^2$$

and

$$2Z/\omega_0\approx(1/Q)(E_{s_1}/V)\cos\phi,$$

giving

$$E^2\approx V^2(\mu M/L-1)^2[1+(E_{s_1}/V)\times\sqrt{1-(2Z/\omega_0)^2(QV/E_{s_1})^2}]^2$$

Substitution in equation 25 gives

$$\{2Z/\omega_0+(3V_2^2/V^2+8V_3^2/V^2)-(\epsilon/Q)\times[1/(\mu M/L-1)][1+m-(V^2/V_0^2)\times(1+(E_{s_1}/V)\sqrt{1-(2Z/\omega_0)^2(QV/E_{s_1})^2})^2]\}^2+\epsilon^2[1-(V^2/V_0^2)(1+(E_{s_1}/V)\times\sqrt{1-(2Z/\omega_0)^2(QV/E_{s_1})^2})^2]^2=(E_{s_1}^2/V^2)\epsilon^2\left[1+m-(V^2/V_0^2)\times\left(1+\frac{E_{s_1}}{V}\sqrt{1-\left(\frac{2Z}{\omega_0}\frac{QV}{E_{s_1}}\right)^2}\right)^2\right]^2\quad(26)$$

Equation 26 gives the ranges of synchronization. Whether the stationary amplitudes given occur in practice or not depends on their stability. The condition of stability can be obtained following the same lines as given in reference 1.

Appendix

Replacing the current i by Q, the differential equation 3 will be the equation of the capacitor charge Q in a series R_t-L-C_t circuit, where R_t and C_t are periodically varying and given by

$$R_t=R_{0t}\left[1+\sum_{m=1}^{\infty}h_m\cos m(\textstyle\int\omega dt+\theta_t)\right]$$

$$C_t=[C/(1+rh_{0t})]/\left[1+\sum_{m=1}^{\infty}k_m\times\cos m(\textstyle\int\omega dt+\theta_t)\right]$$

while the inductance L is constant.

To solve for Q, let $Q=qe^F$ and choose the function of time F such that dq/dt vanishes and neglect terms containing high powers of $R_{0t}/\omega L$ and $R_{0t}h_m/\omega L$. In this way, F will be given by

$$F=-(1/2L)\int R_{0t}\left[1+\sum_{m=1}^{\infty}h_m\times\cos m(\textstyle\int\omega dt+\theta_t)\right]dt$$

and the differential equation becomes

$$di_q/dt+\omega_{0_1}^2\left\{1+(R_{0t}/2\omega L)(\omega^2/\omega_{0_1}^2)\sum_{m=1}^{\infty}\times mg_m\sin\,[m(\textstyle\int\omega dt+\theta_t)+\theta_m]\right\}\int i_q dt=0\quad(27)$$

where

$$i_q=dq/at$$

$$g_m=h_m\sqrt{1+\{[r/(L/C)(1-\mu M/L)]\times(2\omega L/m)(\omega_0^2/\omega^2)\}^2}$$

$$\approx h_m\sqrt{1+\{[r/(L/C)(1-\mu M/L)]\times(2\omega L/m)\}^2}$$

$$\tan\theta_m\approx[r/(L/C)(1-\mu M/L)](2\omega L/m)$$

$\omega_{0_1}^2$, R_{0t}, ω^2, and θ_t are slowly varying functions of time compared with the frequency of oscillation. They could be considered constants for an interval containing several cycles of oscillations. For this small time interval, equation 27 will first be solved.

Let $\omega_{0_1}=\omega(1+d)$ and $d=$amount of detune, and it is assumed to be very small compared with unity. The connection diagram giving the current i_q of equation 27 will be as shown, Fig. 3, where

$$1/C_1=\omega^2L,\ 1/C_2=(2d+d^2)/C_1,\ 1/C_{1t}=(1/C_1)(R_{0t}/2\omega L)\sum_{m=1}^{\infty}mg_m\times\sin\,[m(\omega t+\theta_t)+\theta_m]$$

According to the well-known compensation theorem, if short-circuiting C_2 and C_{1t} by closing the artificial key AK and replacing instead a generator of zero internal impedance and having an electromotive force equal to $-[(1/C_2)+(1/C_{1t})]\int i_q dt$, the actual current i_q flowing in the circuit would be unchanged; but i_q is unknown; therefore successive approximations will be adopted as follows: Closing the AK, the first approximation of i_q would be given by $i_1=A\sin(\omega t+\phi)$, where A and ϕ are constants to be determined from initial conditions. For a second approximation, C_2 and C_{1t} are replaced by a generator of zero internal impedance and having an electromotive force $=v_1=-[(1/C_2)+(1/C_{1t})]\int i_1 dt$. Closing AK of Fig. 3 and applying the component of v_1 of frequency $\omega/2\pi$, the current will be given by

$$i_2\approx A[(2d+d^2)/2]\omega t\cos(\omega t+\phi)+(R_{0t}/\omega L)(g_2\omega/4)t\sin(\omega t+2\theta_t+\theta_2-\phi)$$

The fundamental component of iq is given by (i_1+i_2) to a second degree of approximation. Following the same lines, other approximations can be obtained:

$$i_3=A\{[(R_{0t}/\omega L)(g_2\omega/4)]^2-[(2d+d^2)(\omega/2)]^2\}(t^2/2!)\sin(\omega t+\phi)$$

$$i_4=A\{[(R_{0t}/\omega L)(g_2\omega/4)]^2-[(2d+d^2)(\omega/2)]^2\}(t^3/3!)\times[(2d+d^2)(\omega/2)\cos(\omega t+\phi)+(R_{0t}/\omega L)(g_2\omega/2)\sin(\omega t+2\theta_t+\text{[illegible]}-\phi)]$$

$$i_5=A\{[(R_{0t}/\omega L)(g_2\omega/4)]^2-[(2d+d^2)\times(\omega/2)]^2\}^2(t^4/4!)\sin(\omega t+\phi)$$

Therefore, to the first power of R_{0t}, $R_{0t}g_m$, and $(2d+d^2)$, the fundamental of the current i_q is given by

$$i_{q_{\text{fund}}}=i_1+i_2+i_3+i_4+i_5+\ldots=B_1\epsilon^{\sqrt{a}t}\sin(\omega t+\phi')+B_2\epsilon^{-\sqrt{a}t}\sin(\omega t+\phi'')\quad(28)$$

where

$$\sqrt{a}=(\omega/4)\sqrt{(g_2R_{0t}/\omega L)^2-[2(2d+d^2)]^2}$$

$$\tan\phi'=[(g_2R_{0t}/\omega L)\sin(2\theta_t+\theta_2)+2(2d+d^2)]/[(4\sqrt{a}/\omega)+(g_2R_{0t}/\omega L)\cos(2\theta_t+\theta_2)]$$

$$\tan\phi''=-[(g_2R_{0t}/\omega L)\sin(2\theta_t+\theta_2)+2(2d+d^2)]/[(4\sqrt{a}/\omega)-(g_2R_{0t}/\omega L)\cos(2\theta_t+\theta_2)]$$

B_1 and B_2 are constants depending on initial conditions. The term containing $\epsilon^{-\sqrt{a}t}$ in equation 28 will be neglected as it is a highly damped one. The harmonics other than the fundamental are obtained as follows: replace C_2 and C_{1t} of Fig. 3 by a generator of zero internal impedance and having an electromotive force

$$v_h=-[(1/C_2)+(1/C_{1t})]\int i_{q_{\text{fund}}}dt=-B_1\epsilon^{\sqrt{a}t}(R_{0t}/4\sqrt{1+a/\omega^2})\times\Big((4\omega L/R_{0t})(2d+d^2)\sin(\omega t+\phi'-\theta')+g_1\cos\,[\theta_t+\theta_1-(\phi'-\theta')]+2g_2\times\cos\,[\omega t+2\theta_t+\theta_2-(\phi'-\theta')]+\sum_{m=3}^{\infty}\times\{mg_m\cos\,[(m-1)\omega t+m\theta_t+\theta_m-(\phi'-\theta')]-(m-2)g_{m-2}\cos\,[(m-1)\omega t+(m-2)\theta_t+\theta_{m-2}+(\phi'-\theta')]\}\Big)$$

where

$$\tan \theta' = \omega/\sqrt{a}$$

The effect of the fundamental component of v_h of frequency $\omega/2\pi$ is included in the predetermined fundamental, while the effects of other components give the harmonic currents including the zero-frequency one. Therefore, the current i_q approximated to the first power of R_{0t}, $R_{0t}g_m$, and $(2d+d^2)$ will be given by

$$i_q \approx B_1 \epsilon^{\sqrt{a}t}\Big(\sin(\omega t+\phi') - (\omega^2/\omega_{0_1}{}^2) \times (g_m R_{0t}/4\omega L)(\sqrt{a}/\omega)\cos(\theta_t+\theta_1- \phi'+\theta') + (\omega^2/\omega_{0_1}{}^2) \times (R_{0t}/4\omega L) \times \sum_{m=3}^{\infty} \{(m-1)/[(m-1)^2-1]\} \times \{(m-2)g_{m-2}\sin[(m-1)\omega t+(m-2)\theta_t+ (\phi'-\theta')] - mg_m \sin[(m-1)\omega t+ m\theta_t+\theta_m-(\phi'-\theta')]\}\Big) \quad (29)$$

The current i_q given by equation 29 is true for the specified time interval which contains several cycles of oscillations. For a general solution at any time t, $\sqrt{a}\,t$ should be replaced by $\int\sqrt{a}\,dt$ and ωt by $\int \omega\, dt$. The charge Q is thus given by

$$Q = \epsilon^{F}\int i_q dt = (B_1/\omega)\epsilon^{\int[\sqrt{a}-(R_{0t}/2L)]dt} \times \Big(\sin\left(\int\omega dt+\phi'-\theta'\right) - (R_{0t}/4\omega L)\times (\omega^2/\omega_{0_1}{}^2)g_m\cos(\theta_t+\theta_1-\phi'+\theta') - (R_{0t}/4\omega L)(\omega^2/\omega_{0_1}{}^2)\sum_{m=3}^{\infty}\{1/[(m-1)^2- 1]\}\{(m-2)g_{m-2}\cos[(m-1)\textstyle\int\omega dt+ (m-2)\theta_t+\theta_{m-2}+(\phi'-\theta')] - mg_m\times \cos[(m-1)\textstyle\int\omega dt+m\theta_t+\theta_m- (\phi'-\theta')]\} - (R_{0t}/4\omega L)\sum_{m=1}^{\infty}(h_m/m)\times \{\cos[(m-1)\textstyle\int\omega dt+m\theta_t-(\phi'-\theta')] - \cos[(m+1)\textstyle\int\omega dt+m\theta_t+(\phi'-\theta')]\}\Big) \quad (30)$$

Equation 30 gives also the current i of equation 3. However, the fundamental of the current i must be of the same form as assumed and which is $i = I_t \cos \int\omega dt$. Therefore a possible solution for equation 3 may be obtained from equation 30 provided $(\phi'-\theta') = \pi/2$. Putting $(\phi'-\theta') = \pi/2$ in equation 30, equation 11 in the manuscript is obtained. Substituting ϕ' and θ' by their expressions, neglecting second order effects, and rearranging result also in equation 12.

References

1. Effects of Harmonics on the Frequency of Oscillation as Well as on the Asymmetry of the Resonance Curves, Abd El-Samie Mostafa. *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 309–14.

2. The Application of Power Series to the Solution of Non-Linear Circuits, A. W. Gillies. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 96, 1949, p. 453.

Very-High-Frequency Radio Link Between Puerto Rico and the Virgin Islands

ROGER McSWEENY

ASSOCIATE MEMBER AIEE

PUERTO Rico as well as Cuba, Hispaniola, and Jamaica comprise the Greater Antilles or West Indies, Puerto Rico being the most easterly of the group. About 50 to 75 miles farther east of Puerto Rico are the Virgin Islands which are at the end of an extensive chain of islands called the Lesser Antilles. This paper is concerned with Puerto Rico and with Saint Thomas and Saint Croix of the Virgin Islands, which are about 18 degrees north of the equator, 1,000 miles east of Miami, Fla., and 1,600 miles southeast of New York City. The Virgin Islands are closely associated through both travel and trade with Puerto Rico and, consequently, require intercommunication with it and through it to the American mainland. Puerto Rico is a commonwealth which is associated with the United States while Saint Thomas and Saint Croix are territories of the United States.

Prior to installation of the very-high-frequency (vhf) links to be described in this paper, telecommunication facilities available for public use were limited to submarine telegraph cables and to high-frequency radiotelegraph and radiotelephone circuits. The latter were not very satisfactory because of unstable sky-wave transmission over distances of about 70 miles and the high levels of noise and interference on amplitude-modulated frequencies between 4 and 7 megacycles (mc).

The new facilities, operated by the Radio Corporation of Porto Rico and the All American Cables and Radio, Inc., are based on a 150-mc frequency-modulated system which provides several telephone channels that connect with public telephone systems of the islands. In addition, telegraph printer channels are provided to cable offices for general public use and other printer channels are leased to airlines and shipping companies. The telephone services are extended to the mainland through high-frequency radio circuits from Puerto Rico to the United States. Telegraph services are extended to the mainland through the cable or radio facilities of All America Cables and Radio, Inc.

General Description

Fig. 1 is a map of the system. There are two vhf radio links involved, one of 69 miles from Sabana Llana, Puerto Rico, to Crown Mountain, Saint Thomas, and one of 45 miles from Crown Mountain to Christiansted on Saint Croix. At the Puerto Rican end, the Sabana Llana radio station is connected through 10 miles of cable to the telephone exchange and telegraph office in the city of San Juan. Both radio equipment and carrier apparatus for deriving the telephone channels are located at Sabana Llana.

On Saint Thomas, the Crown Mountain radio station is connected through 4 miles of cable to the town of Charlotte Amalie where carrier equipment, telephone exchange, and telegraph terminals are located. On Saint Croix, the radio and terminating equipments are both at the same location in the town of Christiansted.

Voice and carrier provide four telephone channels on the radio link from Sabana Llana to Crown Mountain and this may be increased to five channels if required. Of the present four channels

---

Paper 55-751, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted March 2, 1955; made available for printing July 28, 1955.

Roger McSweeny is with the American Cable and Radio Corporation, New York, N. Y.

The successful operation of the system has been due to the work of the author's associates, K. M. Barbier at Saint Thomas, and P. A. Girard and J. Castanera in Puerto Rico.

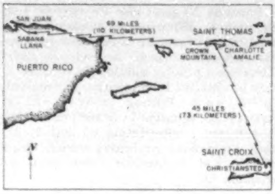


Fig. 1. Map of system

the first is transferred at Crown Mountain to the Christiansted link to provide a direct telephone circuit between Saint Croix and San Juan. The second is connected at Sabana Llana to other radio equipment to provide a direct telephone circuit between Charlotte Amalie and New York. The third is for telephony between Charlotte Amalie and San Juan, and the fourth is subdivided by voice-frequency carrier equipment into four telegraph printer channels. This circuit may be increased to nine printer channels if needed.

Voice and carrier provide three telephone channels on the radio link from Crown Mountain to Christiansted and this may be increased to four channels if required. Of the present three channels, the first is transferred at Crown Mountain to San Juan; the second provides a telephone circuit between Charlotte Amalie and Christiansted and the third now carries a telegraph printer channel. This may be increased to four printer channels, if needed. The system may be expanded in the future to connect two other Virgin Islands, Saint John and Tortola, to Saint Thomas through the Crown Mountain radio station.

Frequencies and Sites

The radio frequencies employed are listed in Table I. While lower vhf frequencies probably would produce better results over these paths, the frequencies used are the lowest for which licenses could be obtained.

Prior to selecting the station site on Saint Thomas, a small transmitter was set up at Sabana Llana and operated for six weeks on 158.01 mc with a corner reflector antenna mounted 260 feet above ground. Received field strength was observed at 19 locations on Saint Thomas and continuous recordings of 1 week each were made at 6 locations. These tests indicated that a substantially stronger and steadier signal could be obtained at the top of Crown Mountain than at any other location on the island.

As a result of this rather remote station location, the cost of radio equipment at Crown Mountain was only a fraction of the cost of the necessary auxiliary facilities such as improvements to roadway, construction of power and telephone lines, building, etc.

Table I. Radio Frequencies

Frequency, Mc	From	To
158.01	Sabana Llana	Crown Mountain
152.75	Crown Mountain	Sabana Llana
152.57	Crown Mountain	Christiansted
157.83	Christiansted	Crown Mountain

Radio Paths

Fig. 2 shows profiles of the two radio frequency paths. The elevations shown are the heights of the antennas above mean sea level. Antenna elevation of 1,580 feet at Crown Mountain is largely responsible for the satisfactory performance of the system. From this elevation, the 45-mile path to Christiansted is above the geometric horizon and presents little difficulty although a first Fresnal zone clearance is not obtained. The 69-mile path to Sabana Llana falls below the smooth-earth geometric horizon but is above the effective radio horizon based on a 4/3 radius earth.

Tests indicated that the maximum practical height of the antenna (290 feet above sea level) using an existing tower at Sabana Llana was 3 and 10 decibels (db) better respectively than the 120-foot and 70-foot heights at the same location. The maximum height does not clear the tops of two hills which project into the path, as shown in Fig. 2.

Measured space losses, excluding estimated antenna gains and line losses, on on the paths from Crown Mountain to Christiansted and Sabana Llana are approximately 110 and 120 db respectively. Calculated free space losses for the same distances between isotropic antennas are 113 db and 117 db. The low loss on the Christiansted path appears to indicate that signals reflected from the surface of the sea between the islands arrive approximately in phase with the directly propagated signals.

Usually there is little fading on either link, but slow fluctuations in field intensity of the order of 3 db occur continuously. In the first 18 months of commercial operation that started in August 1953, several periods lasting from 5 to 15 minutes have been observed during which field intensity dropped to approximately 15 db less than normal in both directions on the Sabana Llana-Crown Mountain link. Commercial operation continued without interruption during these periods. No correlation of the abnormally low fields with weather or time of day has been observed.

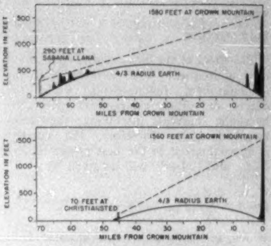


Fig. 2. Radio path profiles

Transmitters and Receivers

The frequency-modulated transmitters employed on the Christiansted-Crown Mountain link have a power output of 25 watts. Similar transmitters followed by power amplifiers of 100-watt output are employed on the Crown Mountain-Sabana Llana link. Peak deviation is limited by regulations of the Federal Communications Commission to ±15 kc for this band. Limiting amplifiers are installed ahead of the transmitters to prevent excessive deviation caused by unusual line levels. The outstanding fea-

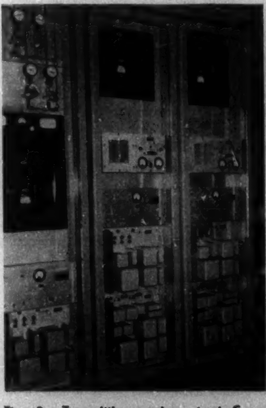


Fig. 3. Transmitting equipment at Crown Mountain

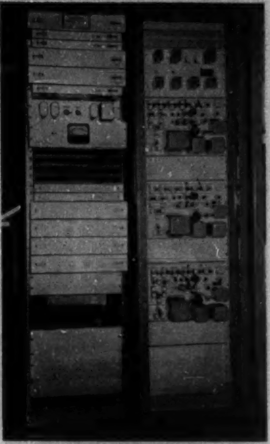


Fig. 4. Line and receiving equipment at Crown Mountain

ol. 74
. 1
55

ture of the transmitters is their low level of distortion.

Fig. 3 shows typical transmitting equipment. The cabinet at left which is partially cut off in the figure contains air-pressure indicators for coaxial lines, a frequency and deviation monitor, a line-limiting amplifier, and a transmitter with power supply. The other two cabinets each contain a line-limiting amplifier, a transmitter with power supply, and a power amplifier with power supply.

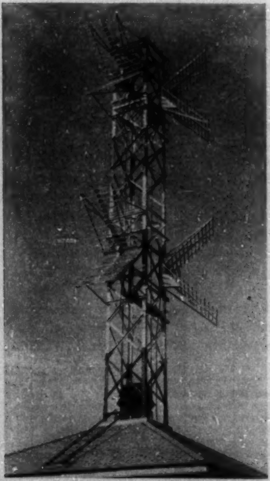


Fig. 5. Antennas at Crown Mountain

The receivers are crystal-controlled double-conversion superheterodynes with automatic volume control, two limiting stages, and a discriminator. The frequency-modulated improvement threshold is 1 microvolt, and the noise figure is 7 db. Signal normally present at the receiver input terminals is approximately 500 microvolts on the Sabana Llana-Crown Mountain link and approximately 1,000 microvolts on the Crown Mountain-Christiansted link.

Fig. 4 shows the typical receiving equipment. The rack at the left contains apparatus associated with the cable lines such as repeating coils, volume indicator, filters, and jacks. The cabinet at the right contains a remote-control relay panel and three receivers.

The transmitter and receiver together are capable of a system response-frequency characteristic that is flat to within 1 db from 200 cycles per second to 20 kc with less than 1-per-cent amplitude distortion and a 50-db signal-to-noise ratio per 3-kc channel. Intermodulation between two test tones is 50 db down. This performance is considerably better than that of frequency-modulated radio equipment commonly used for mobile services or for single-channel telephone service. The higher standards are necessary to prevent crosstalk on multichannel links.

Spare radio equipment is provided at each location but no provision is made for automatic switchover in event of failure. The Crown Mountain station is the only unattended location and, for this point, a monitoring receiver is provided at the office in Charlotte Amalie, which is 4 miles away, with relays for switching transmitters and receivers through the interconnecting cable. Alternating current supplied by local public utilities is used at all locations, but voltage regulators and gasoline-engine emergency generators with automatic starting and transfer facilities are provided.

Antennas

Horizontally polarized corner reflector type of antennas are used. They have a gain of 8 db over a half-wave dipole. With the exception of those at Sabana Llana, all antennas are connected to the equipment through 7/8-inch-diameter air-spaced copper-tube coaxial lines which are fed with dry air at low pressure. Antennas for diversity reception have not been provided because there appears to be no need for them.

Fig. 5 is a view of four antennas, two transmitting and two receiving, mounted on a tower above the station building at Crown Mountain. After this picture was taken, two of the antennas were moved from the tower to poles about 60 feet away to reduce coupling between the Christiansted and Sabana Llana links at Crown Mountain.

Fig. 6 shows the single combined transmitting and receiving antenna on a 300-foot tower at Sabana Llana. Since this antenna is 800 feet from the equipment, $1^3/_8$-inch-diameter coaxial line is employed to hold transmission line losses to 1.6 db. At the station end of the line, cavity resonators, shown in Fig. 7, are used to keep the transmitter radio frequency energy out of the receiver.

At Christiansted the two antennas are mounted on a 100-foot tower behind the telephone building in the center of the town where they are exposed to occasional strong ignition radiation from old automobiles on adjacent streets. Interference from this source is generally maintained at 30 db or better below speech level by the fairly strong frequency-modulated received signal strength and by the horizontally polarized antennas.

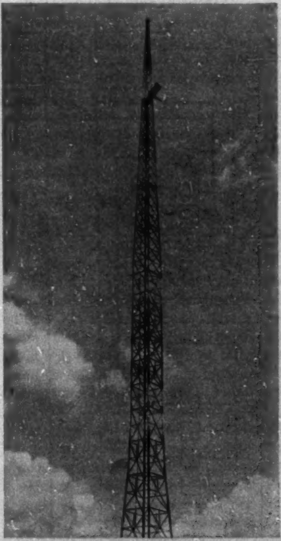


Fig. 6. Antenna and portion of supporting tower at Sabana Llana

Fig. 7. Cavity resonators at Sabana Llana

The stations are in the area where hurricanes originate. The self-supporting towers are designed to withstand safely wind velocities of 120 miles per hour and they are much heavier than towers commonly used for similar service in the United States. The Crown Mountain tower, in the most exposed location, has been strongly guyed in addition.

Carrier Equipment

Fig. 8 is a block schematic of the system. Channel 1 is on voice frequencies while channels 2, 3, and 4 use low-frequency carriers which are of the same frequency in both directions of transmission. Standard telephone-type single-side-band carrier-suppressed equipment is used.

Channel 1 is transferred at Crown Mountain from one link to the other by means of low- and high-pass filters, as indicated in Fig. 8, to provide the direct circuit between Saint Croix and Puerto Rico. Hybrid coils are provided where the 4-wire radio circuits and 2-wire telephone exchanges are interconnected. Two frequency (1,000 cycles per second interrupted 20 times per second) ringers are provided for signaling between telephone operators.

Standard telegraph carrier techniques are used to subdivide one telephone channel into printer channels. Frequency shift-keying is employed and a channel spacing of 120 cycles per second is used. Fig. 9 is a view of the telegraph carrier equipment and monitor printer at one location.

Operating Results

Over-all unweighted signal-to-noise ratio at the telephone switchboards, which includes noise contributions from lines, carrier equipment, radio apparatus, and crosstalk, is about 35 db.

During the initial 18 months of operation, service has been available 99 per cent of the time between San Juan and

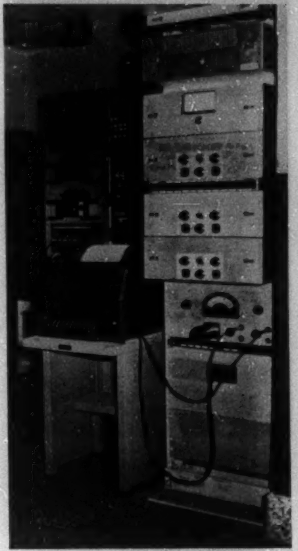


Fig. 9. Telegraph carrier equipment at San Juan

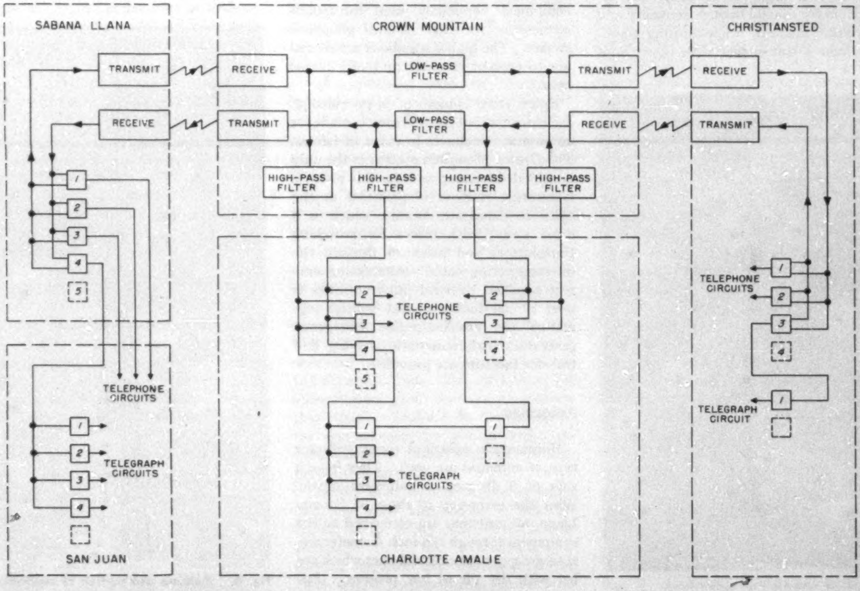

SABANA LLANA
CROWN MOUNTAIN
CHRISTIANSTED
TRANSMIT
RECEIVE
RECEIVE
LOW-PASS FILTER
TRANSMIT
RECEIVE
RECEIVE
TRANSMIT
LOW-PASS FILTER
RECEIVE
TRANSMIT
HIGH-PASS FILTER
HIGH-PASS FILTER
HIGH-PASS FILTER
HIGH-PASS FILTER
1
2
3
4
5
TELEPHONE CIRCUITS
TELEGRAPH CIRCUITS
TELEGRAPH CIRCUIT
SAN JUAN
CHARLOTTE AMALIE

Fig. 8. Block schematic

ol. 74
. 1
955

Charlotte Amalie, and 97 per cent of the time between San Juan and Christiansted. The latter service involves two vhf links in series.

Principal causes of lost time have been defective vacuum tubes, open fuses resulting from power surges, faults in land lines, power failures, and noise caused by poor bonding on antennas and towers. Many of these faults have been partially or wholly corrected and less lost time is expected in the future.

What Is a Minimum-Phase Network?

NORMAN BALABANIAN
NONMEMBER AIEE

WILBUR R. LEPAGE
MEMBER AIEE

THE terms "minimum-phase" and "nonminimum-phase" as applied to network functions were probably first introduced and defined by Bode.[1] Since their introduction these terms have become almost household words in circuit theory circles. There seems to be unanimity of opinion as to the definitions of minimum and nonminimum-phase networks, but there is less than unanimity in understanding of the properties of these two classes of networks. What is it that is minimum about a minimum-phase network? Is it the algebraic value or the absolute value of the phase angle that is a minimum? Is it the total change in phase angle as frequency varies from zero to infinity that is a minimum? Should the angle considered be angle of lead or angle of lag?

Many persons have answered these questions satisfactorily for themselves, and many teachers have supplied satisfactory answers to their students. However, in the opinion of the authors evidence exists that there is still need of clarification; and it is the purpose of this paper to clear up some of the mystery.

General Network Transfer Function

The distinction between minimum and nonminimum-phase functions applies only to transfer functions; functions which represent the ratio of an input to an output, or of an output to an input. The function will be defined as the ratio of the output to the input. Specifically, a 2-terminal pair which is linear, lumped, stable, and with the input labeled "1" and the output labeled "2" will be considered. Thus, the subject will be any one of the steady-state functions E_2/I_1, I_2/E_1, E_2/E_1 or I_2/I_1 which are functions of the complex frequency $s=\sigma+j\omega$. The symbol $F(s)$ will be used to represent any one of these functions. On the $j\omega$ axis this function will be written $F(j\omega)$ and it is understood to be the usual steady-state response function.

The general response function to be considered is a rational function of s which may be written

$$F(s)=Ks^q\times \frac{(s-s_1)\ldots(s-s_r)(s-s_{r+1})\ldots(s-s_l)}{(s-s_1')\ldots(s-s_n')} \tag{1}$$

From standard network theory it is known that there are restrictions on the locations of zeros and poles of this function. Complex zeros or poles occur in complex conjugate pairs, and since the network is stable it is known that no poles can appear in the right half of the s-plane. However, since this is a transfer function, zeros can appear in either half of the s-plane. A multiple zero is possible at the origin, and this possibility is explicitly taken care of by the factor s^q.

The degenerate case of a simple pole at the origin can be included by permitting q to have the value -1. A simple pole at infinity will require the degree of the numerator to be greater than that of the denominator by unity. Thus, all possible cases will be included if the following conditions are specified

$$\begin{aligned} &-1\leqq q\\ &l+q\leqq n+1 \end{aligned} \tag{2}$$

Equation 1 represents a general case in which $s_1, s_2\ldots s_r$ are in the right half of the s-plane. Consider the angle function of equation 1 as s moves through zero along the $j\omega$ axis, in the direction of increasing ω. The angle of the factor s^q goes discontinuously from $-q(\pi/2)$ to $+q(\pi/2)$, and may be defined as being zero at $s=0$. This is illustrated in Fig. 1. In the left half s-plane each factor of the numerator or denominator due to a real critical point (zero or pole) has zero angle at $s=0$. Also, each pair of factors in numerator or denominator due to a pair of complex conjugate critical points has zero angle at $s=0$. Now, consider a factor $(s-s_i)$ due to a real zero in the right half-plane. As s goes through zero along the $j\omega$ axis the angle of this factor goes through π, as shown in Fig. 2(A). If there is a pair of factors due to a pair of complex zeros in the right half-plane, as shown in Fig. 2(B) the angle due to this pair of factors goes through $\pi+\alpha+(\pi-\alpha)=2\pi$ as s goes through zero along the $j\omega$ axis. Thus, each factor due to a zero on the right of the s-plane contributes an angle π at zero frequency. This will mean that the angle function will not be an odd function of frequency. (It is possible to make the angle function defined in Fig. 2 odd by introducing a discontinuity of an appropriate multiple of 2π at the origin. However, this is not done because such a discontinuity would amount to jumping from one Riemann surface to another when passing through the origin, and it would invalidate the results about to be derived.)

Now, write $F(s)$ so that its angle function will be an odd function of frequency. To do this, invert the order of each of the factors due to right half s-plane zeros, writing (s_i-s) for $i=1,2\ldots r$. Fig. 3 illustrates that these factors now have angles which go through zero at the origin. This inversion is compensated for by introducing a factor $(-1)^r$ and obtaining

$$F(s)=K(-1)^r s^q\times \frac{(s_1-s)\ldots(s_r-s)(s-s_{r+1})\ldots(s-s_l)}{(s-s_1')\ldots(s-s_n')} \tag{3}$$

The function is now in such a form that the angles of s^q and of the rational function to its right both pass through zero. However, if K is positive and r is odd there will still be an angle of π at the origin, and therefore $F(s)$ will not have an odd angle function of frequency. This trouble can be avoided by stipulating that K shall be negative if r is odd. This condition has significance only when dealing with a specific example, and amounts to stipulating what reference conditions shall be used. For example, if $F(s)$ were to be E_2/E_1 of Fig. 4 it would be necessary to choose the reference polarities as shown in order for $K(-1)^r$ to be positive (i.e.,

---

Paper 55-707, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3-7, 1955. Manuscript submitted June 8, 1955; made available for printing August 15, 1955.

NORMAN BALABANIAN and WILBUR R. LEPAGE are with Syracuse University, Syracuse, N. Y.

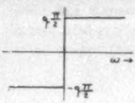


Fig. 1. Angle of s^q when $s=j\omega$ and ω passes through zero

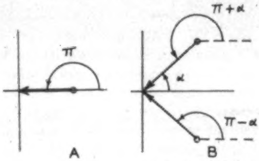


Fig. 2. Angles at $s=0$ of factors $(s-s_1)$ and $(s-s_1)(s-s_1^*)$ where s_1 is in the right half s-plane

for K to be negative). Assuming K is defined in this way for every example to which this general case is to be applied, we can finally write

$$F(s)=|K|s^q\times \frac{(s_1-s)\ldots(s_r-s)(s-s_{r+1})\ldots(s-s_l)}{(s-s_1')\ldots(s-s_n')} \tag{4}$$

We are really interested in the angle of this function when $s=j\omega$, and so we consider

$$F(j\omega)=|K|(j\omega)^q\times \frac{(s_1-j\omega)\ldots(s_r-j\omega)(j\omega-s_{r+1})\ldots(j\omega-s_l)}{(j\omega-s_1')\ldots(j\omega-s_n')} \tag{5}$$

$$=|F(j\omega)|e^{-j\theta(j\omega)} \tag{6}$$

Special note should be taken of the negative sign associated with the angle in equation 6. $\theta(j\omega)$ is defined as the angle by which the output lags the input. The angle θ as a function of ω for various possible locations of the zeros of $F(s)$ will be examined.

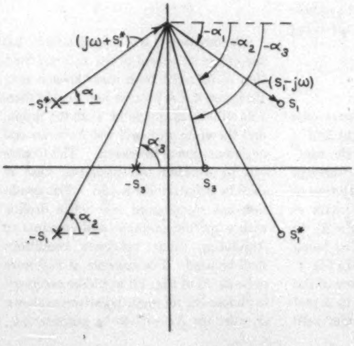


Fig. 5 (left). Angles due to zeros and poles of all-pass function

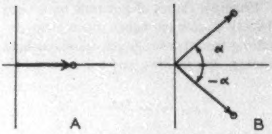


Fig. 3. Angles at $s=0$ of factors (s_1-s) and $(s_1-s)(s_1^*-s)$ where s_1 is in the right half s-plane

Minimum and Nonminimum Phase Functions

In equation 5 let s_1 and $s_2=s_1^*$ be a pair of complex conjugate zeros and s_3 a real zero, these being typical zeros in the right half-plane. $F(j\omega)$ can be written in the modified form

$$F(j\omega)=F_k(j\omega)A(j\omega) \tag{7}$$

where

$$F_k(j\omega)=|K|(j\omega)^q\times \frac{(j\omega+s_1)(j\omega+s_1^*)(j\omega+s_3)(s_4-j\omega)\ldots(s_r-j\omega)(j\omega-s_{r+1})\ldots(j\omega-s_l)}{(j\omega-s_1')\ldots(j\omega-s_n')} \tag{8}$$

$$A(j\omega)=\frac{(s_1-j\omega)(s_1^*-j\omega)(s_3-j\omega)}{(j\omega+s_1)(j\omega+s_1^*)(j\omega+s_3)} \tag{9}$$

The only difference between F_k and F is that $K=3$ of the zeros of F that were in the right half-plane have been replaced in F_k by zeros in the left half-plane obtained by reflecting the three right half-plane zeros in the $j\omega$ axis. Observe that each of the factors in the denominator of equation 9 is the conjugate of one of the factors in the numerator; it is then concluded that

$$|A(j\omega)|=1 \tag{10}$$

From equation 7 is then follows that

$$|F_k(j\omega)|=|F(j\omega)| \tag{11}$$

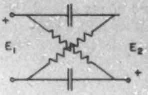


Fig. 4. Example of choice of reference conditions for continuous, odd-angle function of frequency

Now, write

$$F_k(j\omega)=|F(j\omega)|e^{-j\theta_k} \tag{12}$$

$$A(j\omega)=e^{-j\phi} \tag{13}$$

From equation 7 the angles are related by

$$\theta=\theta_k+\phi \tag{14}$$

The function $A(j\omega)$ is called an "all pass function" because its absolute value is 1.

It will now be shown that ϕ is always positive when $\omega>0$ The poles and zeros of $A(s)$ together with the angles contributed by each are shown in Fig 5. From the figure we see that the angle is given by

$$\begin{aligned}\phi=&-(\text{sum of angles due to zeros}-\text{sum of angles due to poles})\\ &=\alpha_1+\alpha_2+\alpha_3-(-\alpha_1-\alpha_2-\alpha_3)=2(\alpha_1+\alpha_2+\alpha_3)\end{aligned} \tag{15}$$

Note that each of the α's is always positive, and therefore that ϕ is positive. Going back to equation 14 the result now appears

$$\theta_k(j\omega)<\theta(j\omega) \tag{16}$$

with emphasis on the fact that this relationship is valid at all values of $\omega>0$. This discussion pertains to a typical case of a pair of complex zeros and one real zero. A similar result would be obtained if only one or only a pair of zeros had been shifted, or if more of them had been shifted. Thus, starting with the original function F, zeros could be shifted, one or a pair at a time, leading to sequence of functions, of which F_k is typical, as follows

$$\begin{aligned}F_1&=|F(j\omega)|e^{-j\theta_1(j\omega)}\\ F_2&=|F(j\omega)|e^{-j\theta_2(j\omega)}\\ &\ldots\ldots\\ F_m&=|F(j\omega)|e^{-j\theta m(j\omega)}\end{aligned} \tag{17}$$

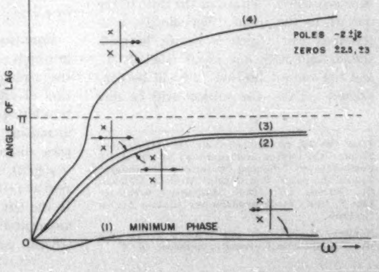


Fig. 6 (right). Illustration of minimum phase and nonminimum phase functions having the properties

$$\theta_m < \ldots < \theta_2 < \theta_1 < \theta \qquad (18)$$

at all values of $\omega > 0$. F_m is the final function when all zeros have been shifted to the left half s-plane.

The usual definition of a minimum-phase function is that it has no zeros in the right half-plane. F_m is therefore a minimum-phase function by this definition. It may be concluded therefore that at all frequencies the angle of lag (θ_m) of a minimum-phase function has the smallest value of all the possible functions having the same magnitude function $|F(j\omega)|$.

Total Change of Phase

The interpretation given by most writers seems to be that a minimum-phase network has the smallest possible change in the angle of lag as the frequency goes from zero to infinity. Let this change be

$$\Delta\theta_m = \theta_m(j\infty) - \theta_m(0+) \qquad (19)$$

where the 0+ signifies the angle just to the positive side of zero on the ω axis. We have seen that due to the factor s^q the angle of lag (the negative of the angle shown in Fig. 1) is $-q\pi/2$ at $\omega = 0+$. All other factors in equation 4 have zero angle at 0+ and thus

$$\Delta\theta_m = \theta_m(j\infty) + \frac{q\pi}{2} \qquad (20)$$

At infinite frequency the angle is $\pi/2$ times the order of the zero at infinity (the degree by which the denominator of equation 4 exceeds the numerator). Thus

$$\theta_m(j\infty) = (n-l-q)\frac{\pi}{2} \qquad (21)$$

and

$$\Delta\theta_m = (n-l-q)\frac{\pi}{2} + q\frac{\pi}{2} \qquad (22A)$$

$$= (n-l)\frac{\pi}{2} \qquad (22B)$$

Exclude for the moment the degenerate cases of poles at zero or infinite frequency, which means that relations 2 become

$$\begin{aligned} &0 \leqq q \\ &l + q \leqq n \end{aligned} \qquad (23)$$

Then in equation 22(A) $n-l-q$ and q are each positive and are, respectively the multiplicities of zeros at $s = \infty$ and at $s = 0$. This result is in agreement with a statement given by Bode.[1]

Now, return to the general case. Since equation 18 is valid at all frequencies (and specifically for $s = j\infty$) and since $\theta(0+)$ will be the same for all functions $F_1, F_2 \ldots F_m$, a relationship similar to equation 18 can also be written for the changes in angle as ω goes from zero to infinity.

$$\Delta\theta_m < \ldots < \Delta\theta_2 < \Delta\theta_1 < \Delta\theta \qquad (24)$$

This last relationship says that the change in angle of lag as ω goes from zero to infinity is algebraically least for the minimum-phase function for all cases whether or not there is a pole at $s = 0$ or $s = \infty$. It includes the possibility that $\Delta\theta_m$ may be negative in the degenerate cases.

Another interesting interpretation is obtained by looking at equation 15. At $s = j\infty$, $\alpha_1 + \alpha_2 + \alpha_3 = 3\pi/2$, and so at $j\infty$ the total reduction in angle is 3π when three zeros are shifted from right to left. This result may be generalized by saying that every time a zero is shifted from the right to the left half s-plane $\theta(j\infty)$ and therefore $\Delta\theta$ is reduced by π. Therefore if p zeros are shifted it will reduce $\Delta\theta$ by $p\pi$. More explicitly, if p zeros are shifted to the left in going from θ_k to θ_{k+1}, then

$$\Delta\theta_k = \Delta\theta_{k+1} + p\pi \qquad (25)$$

Illustrative Example

An example is shown in Fig. 6 in which the positions of two real zeros are permuted. The important properties to be observed are that curve 1, the minimum-phase case, is lower than all the others, and that case 4, with all zeros in the right half-plane, has the highest curve at all frequencies. Equations 18, 24, and 25 are demonstrated by this example.

Conclusion

In network theory it is customary to define minimum and nonminimum-phase networks in terms of positions of zeros of their steady-state response functions. Network specialists have been aware for many years that, for a given amplitude response function, minimum phase networks have the smallest possible change in phase as frequency goes from zero to infinity. However, by defining the network function in such a way that the angle passes continuously through zero at zero frequency (except if there is a zero at the origin) it has been shown here that the angle function of a minimum-phase network is minimum in another sense. Specifically, it has been shown that if the angle is defined as the angle by which the output lags the input, then at all frequencies the angle of the minimum phase function is smaller than the corresponding nonminimum-phase function. It is important to note that this result depends on defining the response function and its angle properly.

Reference

1. Network Analysis and Feedback Amplifier Design (book), H. W. Bode. D. Van Nostrand & Company, Inc., New York, N. Y., 1945, p. 242.

◆

Discussion

T. J. Higgins (University of Wisconsin, Madison, Wis.): This paper is a most instructive one. As one whose current prime interest in networks is in connection with the teaching of feedback control systems theory, I tried to summarize the main points for subsequent class use. The following would appear to encompass my summary of the salient points. First, the authors demark the type of network considered: namely, those whose transfer functions have no poles in the right-hand half-plane (per their constraint as to stable networks); no pairs of conjugate poles on the imaginary axes (at least none such are considered in their analysis); and at most a single pole at the origin (presumably, since a multiple-pole thereat would result in an unstable network). Transfer networks satisfying these conditions plus the additive condition that no zeros exist in the right-hand half-plane are defined as minimum-phase networks.

Second, of such networks the authors show by a well-detailed and clearly set-forth analysis that if the angle of the network function is defined as the angle by the output lags the input and that if the network function is so defined that this angle passes continuously through zero at zero frequency (except if there is a zero at the origin), then of all functions having the same magnitude function, the minimum-phase function has at all frequencies an angle smaller than that of a nonminimum-phase function, and also has the smallest possible change in phase as frequency goes from zero to infinity.

Third, these facts, as also in equation 25, are well-illustrated by the numerical example pertinent to the curves of Fig. 6.

In the thought that this summary epitomizes the essential restrictions and results of the paper, the discusser would congratulate the authors for their clear elucidation of the subject.

In conclusion, I would raise the following question. Suppose that the definition of a minimum-phase network is as above except that the transfer function of the network may also have pairs of single conjugate imaginary poles on the imaginary s-axis. Such a network may be characterized as limitedly stable, since though such poles will contribute oscillatory terms in the output, these are yet of finite magnitude (much as a single pole at the origin contributes a finite constant term in the output to delta-function input). What are the corresponding minimum properties, if any, of this "relaxed minimum-phase network?"

This question is of practical interest in that the transfer functions of some networks used in feedback control systems work are precisely of this type.

N. Balabanian and **W. R. LePage:** The minimum-phase property of a network function (locations of the zeros of the output over input function) is independent of the locations of the poles. However, unstable functions are not of interest, and so the properties of the angle function are discussed only for stable functions.

The question of how zeros of the function on the $j\omega$ axis affect the minimum-phase property may also be asked. It should be remembered that a comparison is to be made of the properties of the angle function of two (or more) transfer functions, each having the same magnitude function. In the case of a transfer function with zeros on the $j\omega$ axis, there is no other function having the same magnitude with which to compare angles. This is a limiting situation; since there is only one possible angle function, it may be said that this is the minimum. Hence, it may be useful to include in the class of minimum-phase functions those with some or all of their zeros on the $j\omega$ axis.

The Underwater Spark; A Photographic Light Source of High Intrinsic Brilliance

H. C. EARLY
ASSOCIATE MEMBER AIEE

E. A. MARTIN
NONMEMBER AIEE

WHEN an impulse voltage of sufficient magnitude is applied to underwater electrodes, a gaseous conduction channel forms across the gap. If the rate of current rise is thousands of amperes per microsecond, the growth of the gaseous channel is greatly restrained by the inertia of the water and this results in unusually high gas pressure, current density, and temperature. The rapid vaporization of the water at the channel boundary also contributes to the high gas pressure.

The high-current underwater spark has several unique properties which suggest its usefulness as a research tool. Some of these properties are:

1. The gas pressure in the discharge channel is exceptionally high and the amount of light radiated per square millimeter (mm) is unusually large, corresponding to a black body at over 33,000 degrees Kelvin.
2. The duration of the light pulse can be made much shorter than the current pulse by using the "shatter shutter" principle, to be described.
3. There is a high concentration of energy in the ultraviolet which is advantageous in situations where ultraviolet sensitive film can be used.

The information in this paper regarding the light-emitting properties of an underwater spark is a by-product of a more basic investigation of gaseous conduction mechanisms at very high pressures and temperatures. Little effort has been made to optimize the experimental conditions towards the requirements of a photographic light source. The amount of stored energy used in the tests was much greater than necessary for a short-duration light pulse. It is hoped that the information in this paper may indicate some of the potentialities and limitations which would be encountered if further development were undertaken.

Apparatus

Most of the tests to be described were made with a capacitance of 6 microfarads charged to 25 kv. The capacitors were connected to the underwater electrodes with parallel copper straps to reduce inductance. The discharge was initiated by means of a manually operated switch. Because of the intense pressure wave generated by the spark, the steel water tank (of about 2-cubic-foot capacity) was fitted with a splashproof cover, and the Plexiglass observation window was mounted in shock-absorbing rubber. To minimize electrical interference with the measuring equipment, all of the discharge circuitry was enclosed in a shielded metal cabinet.

It was found desirable to control the position of the spark by "exploding" a fine 0.001-inch-diameter wire connected between the electrodes. The use of the wire avoids the erratic spark path which is otherwise obtained, and permits experimental data to be closely reproducible.

Properties of Underwater Sparks

Fig. 1 is an oscilloscope trace of the discharge current when the electrode separation was 1.5 centimeters (cm). The first half-cycle of current lasts approximately 4 microseconds. When the electrode separation is increased to 8 cm, the circuit is approximately critically damped.

Fig. 2 is a sequence of Kerr cell photographs showing the growth of the spark channel. The exposure time for each photograph was 1/4 microsecond. Each picture was taken with a different spark, but with excessively longer time delays before the electro-optical shutter was triggered. At the instant of current maximum the apparent channel diameter was approximately 2.5 mm. The apparent discontinuity in the channel was caused by a small air bubble which was attached to the initiating wire previous to firing

In Fig. 3 a screen was located 1 cm. behind the spark. The light from this screen was bent while passing through the higher density compressed water near the spark channel. At the instant this photograph was taken the expanding cylindrical compression wave in the water had traveled outward about 3 mm. from the spark channel. By measuring the apparent shift in position of the wires on the screen, the change in optical index and hence the density and pressure in the water were calculated. In the test the pressure just inside of the cylindrical shock front was approximately 5,000 atmospheres.

Measurement of Light Output

The usual practice of describing a light source in terms of quantities such as lumens and candles is not applicable for a photographic spark source because these quantities are related to the response curve of the human eye. Since the underwater spark produces far more energy in the ultraviolet than in the visible portion of the spectrum, and since most types of film are especially sensitive in the ultraviolet, it appears relevant to express the brilliance in terms of more fundamental units.

The radiant energy was measured by means of a phototube and an oscilloscope. Farrand optical filters were employed to permit only a known width of spectrum (approximately 120 angstroms half-bandwidth) to enter the phototube. The phototube and filters were calibrated

---

Paper 55-708, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3-7, 1955. Manuscript submitted April 20, 1955; made available for printing August 8, 1955.

H. C. EARLY and E. A. MARTIN are with the University of Michigan, Ann Arbor, Mich.

This work was sponsored by the Office of Ordnance Research. The authors wish to acknowledge the valuable assistance of Professor W. G. Dow and Haldon L. Smith in connection with this research.

Vol.
Pt. 1
1955

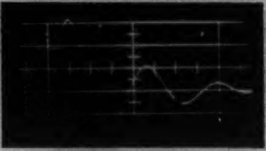


Fig. 1. Oscillogram of discharge current. Two microseconds per division. Peak current, 75,000 amperes

by means of a 108-watt microscope-illuminating lamp (General Electric type-*18A*, 6 volts, 18 amperes). This lamp has a flat filament of tungsten ribbon which is well suited for radiation measurements. It was found necessary to operate the lamp on direct current, to avoid modulation of the light at power frequencies.

Published data regarding the spectral emissivity of tungsten[1] were used to calculate the true filament temperature from pyrometer readings, and also to obtain the intensity of the radiated energy in the passband of the filter. By interposing a mechanical chopper between the filament and the phototube an a-c signal was obtained on the oscilloscope. Phototube sensitivities in terms of microamperes per microwatt, obtained in this manner, agreed closely with the manufacturer's data for the 929 and 926 phototubes used.

The phototubes and circuits were also tested for frequency response against another spark light source of known 2-microsecond duration. The time constant of the phototube circuit was approximately 0.1 microsecond. The effectiveness of the electrical shielding in preventing extraneous pickup was tested by discharging the underwater spark and blocking the light beam. No extraneous pickup could be detected. Fig. 4 is a curve of light output versus time at 4,190 angstroms.

The interpretation of the spectral-energy data is based on the assumption that the spectrum of light from the spark is continuous and has the Planckian distribution characteristic of a black body.[2-4] The high pressures and particle densities, and the high degree of ionization in the underwater spark make the energy transfer between electrons and atoms rapid enough to insure that thermodynamic equilibrium is approached in a small fraction of a microsecond. This subject is completely discussed by Edels.[5] Experiments indicate that the spark channel has a high degree of opacity, and the free path of radiation within the plasma is very short compared to the diameter of the channel, this condition being a prerequisite for black-body emission. A spectrogram of the spark was made with a small Hilger quartz spectrograph. It showed no line structure from 6,900 to 3,400 angstroms, these limits being set by the photographic plate sensitivity and the transmission of the Plexiglass window in the water tank. The assumption of a black-body spectral distribution was further checked by comparing the radiation intensity measured at 4,190 angstroms with the intensity measured at 6,450 angstroms. If the black body assumption is correct, then a temperature calculation based on the intensity at 4,190 angstroms should be in agreement with a temperature calculation based on the data at 6,450 angstroms.

Since the apparent radiating area of the spark channel was available from the Kerr cell photographs, the black-body temperature could be calculated. A possible error in the temperature calculation is due to the uncertainty as to the actual channel diameter. The optical refraction in the compressed water near the channel may magnify the channel in a manner analogous to the magnification of the liquid in a glass thermometer. Thus the calculation gives a lower limit to the source temperature, since the Kerr cell photographs give an upper limit of source size. A temperature calculation for the 25-kv 6-microfarad spark, based on the spectral intensity at 4,190 angstroms and an apparent diameter of 2.5 mm at 2.5 microseconds after initiation gives a source temperature of 33,400 degrees Kelvin. The corresponding calculation based on the measured spectral intensity at 6,450 angstroms indicated a source temperature of 31,100 degrees Kelvin. The discrepancy between these two temperature calculations could well be due to experimental error, rather than to a deviation from the black-body spectral distribution.

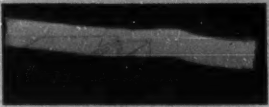


Fig. 3. Kerr cell photograph illustrating refraction of light caused by pressure gradient in water. Forty-three kilovolts, 0.5 microfarad, 2 microseconds after spark initiation

At the instant the apparent channel diameter was 2.5 mm, the measured radiation intensity at 4,190 angstroms, in a direction normal to the spark channel, was 190 watts per millimicron of bandwidth per unit solid angle. The corresponding source brilliance was 1,600 watts per square cm of cylindrical source area per millimicron of bandwidth. The total radiated power from the spark, in the photographically useful range from 3,400 to 7,000 angstroms was 430,000 watts, this being the integrated output in all directions. For special applications where quartz optics could be justified, the more intense radiation below 3,400 angstroms would be available.

Various investigators have reported extreme temperatures in arcs and sparks. Temperatures as high as 50,000 degrees Kelvin are reported in an arc at atmospheric pressure in the center of a whirlpool of water.[6,7] At atmospheric pressure, an arc does not radiate as a black body and therefore the 50,000-degree temperature is not indicative of the light output. At a temperature of 50,000 de-

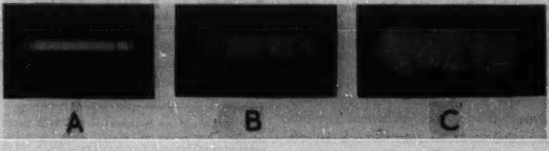


Fig. 2. Sequence of Kerr cell photographs of growth of channel formed by exploding a 1.5-cm length of 0.001-inch-diameter wire under water. Exposure time, 1/4 microsecond

A—One microsecond after initiation of discharge
B—Four microseconds after initiation of discharge
C—Twelve microseconds after initiation of discharge

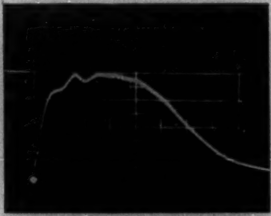


Fig. 4. Phototube measurement of light output. Filter peaked at 4,190 angstroms. Sweep speed, 4 microseconds per division

ol. 74
t. 1
955

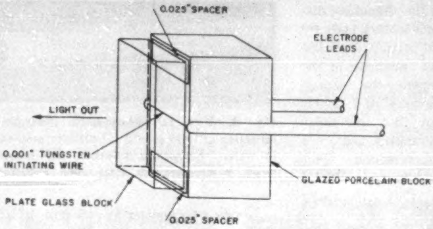


Fig. 5. Schematic drawing of shatter shutter

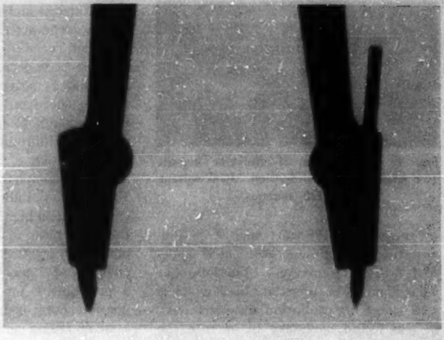


Fig. 7 (right). Shadow picture taken with 1/16-inch-diameter underwater spark light source. Diffraction pattern can be observed at pointed tips

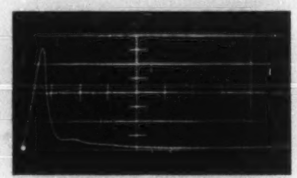


Fig. 6. Oscillogram of light pulse obtained with shatter shutter. Sweep speed, 4 microseconds per division

grees and a pressure of only 1 atmosphere the intensity of radiation is limited by the particle density. Hence an arc at a higher pressure can radiate more light, even though the temperature may be less. Peters[2] reports an arc in a whirlpool of water at 1,000 atmospheres' pressure which radiated as a black body at 12,000 degrees. Glaser[3,4] has measured black-body spark temperatures of 40,000 degrees in argon at 16-atmosphere pressure. The production of high brilliance in a spark light source requires a high gas pressure, and the underliquid spark is an experimentally convenient way of obtaining these pressures.

Shatter Shutter

The light duration of approximately 24 microseconds, indicated in Fig. 4, is excessively long for most applications where spark photography is employed. This duration was greatly reduced by utilizing the steep-front pressure wave in the water to shatter a glass window through which the light passed. There appears to be a variety of ways in which the explosion can be used to open or close optical shutters in time intervals of the order of 1 microsecond. Presumably this technique could be refined to produce intense light pulses of submicrosecond duration.

The shatter shutter arrangement is illustrated in Fig. 5. The 1-mil-diameter tungsten initiating wire was stretched across the surface of the ceramic block which was separated by 0.025-inch shims from the surface of a glass block. This assembly was placed in the water tank with special care to make sure that there were no air bubbles in the space between the glass and ceramic surfaces. After the explosion, the glass was reduced to fine granules. The speed with which the light pulse was attenuated is indicated by Fig. 6, which has the same time base and vertical scale factor as Fig. 4. The extinction of the light was more rapid and complete when the glass block was painted black with the exception of small slits through which the light passed. This black surface absorbed the light which was internally scattered by the crazing of the glass.

Fig. 7 is a shadow picture taken with a 2-microsecond light pulse from a shatter shutter. In this case, the source size was reduced by an aperture having a 1/16-inch-diameter opening, this diameter giving much more light than required to expose the film. In this shadowgraph the light source was 29 feet from the film and the object was 1½ feet from the film. No lens was used.

References

1. The Properties of Tungsten and the Characteristics of Tungsten Lamps, **W. E. Forsythe, A. G. Worthing.** *Astrophysical Journal*, Chicago, Ill., vol. 61, Apr. 1925, p. 146.

2. Temperature and Radiation Measurements on Water Stabilized High Pressure Arcs, **T. Peters.** *Zeitschrift für Physik*, Berlin-Charlottenburg, Germany, vol. 135, Mar. 1953, pp. 573–92.

3. About the Generation of Very High Temperatures in High Power Sparks, **G. Glaser.** *Zeitschrift für Naturforschung*, Tubingen, Germany, vol. 6A, 1951, p. 706.

4. Light Emission from High Current Spark Discharges, **G. Glaser.** *Optik*, Stuttgart, Germany, vol. 7, 1950, pp. 33–53, 61–90.

5. The Determination of the Temperatures of an Electrical Discharge in a Gas, **H. Edels.** *Report No. L/T230*, British Electrical and Allied Industries Research Association, London, England, 1950.

6. Temperature Measurement of Water Stabilized High Power Arcs, **F. Burhorn, H. Maecker, T. Peters.** *Zeitschrift für Physik*, Berlin-Charlottenburg, Germany, vol. 131, Dec. 1951, pp. 28–40.

7. Measuring Plasma Conductivity in the High Power Arc, **H. Maecker, T. Peters.** *Zeitschrift für Physikalische Chemie*, Berlin, Germany, vol. 198, 1951, pp. 318–28.

A Pole-Mounted A-C-Operated Repeater for 12-Channel Open-Wire Carrier Systems

G. W. SEARLE
MEMBER AIEE

R. A. SCHAEFER
ASSOCIATE MEMBER AIEE

ol, 74
t. 1
955

Field of Use

POLE-mounted a-c-operated carrier repeaters have been available since 1952. These repeaters are for use on 4-channel groups and are of the "frequency-frogging" type.[1] When a 12-channel open-wire carrier system using miniature components for both terminals and repeaters became available, it was felt that it might be practical and economical to adapt the repeater to a pole-mounted installation, operated on a-c, with battery stand-by.

It is believed that, for a line with an ultimate requirement of not more than six systems, pole mounting offers economies over a building. Pole-mounted repeaters also allow more freedom in the selection of repeater locations because it should not be necessary to buy or lease land, as would be required for the erection of a building. This freedom of choice in repeater location can eliminate the necessity of retransposing a section of the line to meet the repeater location. As an example, the two repeaters now installed in Wisconsin were located at the poles initially selected. If it had been necessary to purchase or lease land for the erection of a building, it is doubtful whether these locations could have been obtained.

Application

The line chosen for the application of pole-mounted repeaters extends from Madison to Stevens Point, Wisconsin, a distance of about 100 miles; see Fig. 1. The number of pairs of wires on the line varies from 20 pairs in one section to a minimum of 6 pairs in another section.

The conductors are *104* copper with the usual end and pole-pair phantoms. Four pairs on the first arm and two pairs on the third arm can be retransposed for the operation of 12-channel open-wire carrier operating in the 40- to 150-kc range. The entire line is now transposed according to the exposed-line system with alternate-arm transposition for *C* carrier superimposed on the first and third arm. The 12-channel carrier system is operating on one side of the 1 to 4 group. As more systems are added, it will of course be necessary to retranspose the 12-channel carrier pairs.

The Lenkurt *45A* system[2] was selected for this application. The following factors influenced the choice of this type of equipment:

1. It is possible to extend future systems on toll lines where co-ordination with existing *J* systems is required.
2. It is possible to extend future systems by means of cables using *45B* equipment and co-ordinating with existing *N* systems.
3. The repeaters available for the *45A* system use miniature components adaptable to pole mounting.

Repeater Installation

The repeater installation consists of the following essential parts, as shown in Fig. 2. The cabinet on the left contains the repeater, its a-c power unit, and the relays used to transfer the power supply to the stand-by power unit. The center cabinet contains the stand-by power supply. The cabinet on the right contains two protected cable terminals and the crosstalk-suppression filters used on the non-*45A* pairs.

The cabinets used for the repeaters and the suppression filters are *N* repeater cabinets, unmodified except for the provision of the necessary entrance holes for wires and cables. The cabinets are mounted on regular 10-foot crossarms supported on poles, one a regular line pole, the other an additional pole set along the line about 10 feet from the regular pole. All wires are dead-ended, as shown in Fig. 3, and the wire between the poles is removed.

Methods of line treatment at repeater points for open-wire carrier systems operating up to 150 kc have been known for some time. In this installation an attempt was made to determine if the requirements could be relaxed somewhat; first for one system, and subsequently as more systems are added to the line. This may be of particular interest in operating areas where short-haul systems are usually installed. The mounting, bridling, and cabling details are shown in Fig. 4. The carrier line filters are mounted under the crossarm. The low-pass sides of the filters are connected together by means of bridle wire at the top of the pole. The high-pass side of the filters are brought into the repeater cabinet with bridle wire. The non-*45A* pairs are bridled to a protected cable terminal mounted below the crossarms and are brought into the cross-talk-suppression filter cabinet by cable.

The repeater cabinet is shown in Fig. 5. The repeater consists of a regulator, line amplifier, and directional filter for each direction of transmission mounted on a compact shelf. Under normal conditions the repeater is powered from the a-c supply through the power-supply unit shown below the repeater. The power-supply unit contains a filament transformer and a rectifier for plate supply. In case of failure of the a-c supply or failure of the rectifier, the repeater is transferred to the stand-by power unit. The relays for making the transfer are located below the power-supply unit.

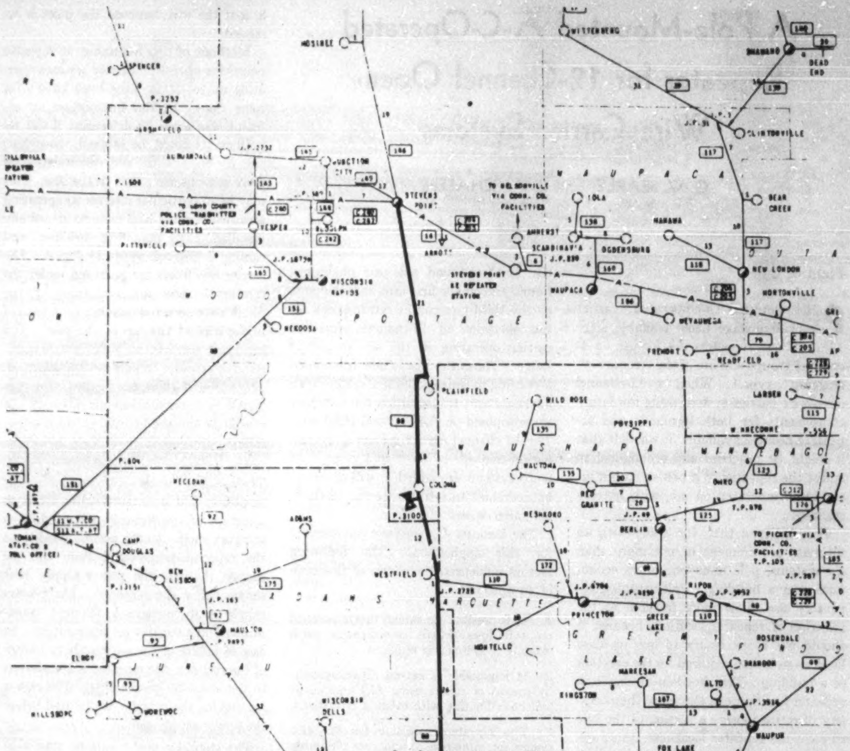
The stand-by power supply, shown in Fig. 6, is the same unit normally used in the type *O* carrier system, with an additional 24-volt battery lead to the transfer relay circuit. It consists of a 50-ampere-hour 24-volt storage battery with its trickle charger and alarm circuits. It also contains a 130-volt dynamotor for plate supply. A time-delay circuit is included to prevent transfer back to the a-c supply during short spurts of power occurring during trouble conditions on the power line. Transfer to the stand-by unit occurs immediately upon failure of the normal a-c supply, but the a-c supply must be reliably restored before the transfer back is made.

An alarm circuit is provided to indicate to the terminal any abnormal conditions of the power supply including transfer to the stand-by power supply. This alarm circuit, used in common with the pilot alarm on the repeaters, is shown on Fig. 7. Tests to locate the alarm conditions are made by means of a Wheatstone bridge. Both the type and location of the ab-

---

Paper 55-682, recommended by the AIEE Wire Communications System Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 3–7, 1955. Manuscript submitted March 24, 1955; made available for printing August 2, 1955.

G. W. SEARLE (deceased) was with the Wisconsin Telephone Company, Milwaukee, Wis., and R. A. SCHAEFER is with the Wisconsin Telephone Company, Milwaukee, Wis.



WITTENBERG
SHAWANO
DEAD END
MOSINEE
SPENCER
MARSHFIELD
AUBURNDALE
JUNCTION CITY
STEVENS POINT
CLINTONVILLE
BEAR CREEK
W A U P A C A
IOLA
AMHERST
SCANDINAVIA
OGDENSBURG
MANAWA
NEW LONDON
NORTONVILLE
WEYAUWEGA
FREMONT
MEDFIELD
ARNOTT
STEVENS POINT RE. REPEATER STATION
WAUPACA
PITTSVILLE
VESPER
RANDOLPH
WISCONSIN RAPIDS
NEKOOSA
P O R T A G E
W O O D
PLAINFIELD
WILD ROSE
W A U S H A R A
WAUTOMA
REDGRANITE
POYSIPPI
W. WINNECONNE
OMRO
LARSEN
W I N N E B A G O
BERLIN
COLOMA
NECEDAH
ADAMS
CAMP DOUGLAS
NEW LISBON
TOMAH
MAUSTON
WESTFIELD
A D A M S
NESHKORO
PRINCETON
GREEN LAKE
RIPON
ROSENDALE
M A R Q U E T T E
MONTELLO
G R E E N   L A K E
MARKESAN
KINGSTON
ALTO
BRANDON
WAUPUN
F O N D
ELROY
J U N E A U
HILLSBORO
WONEWOC
WISCONSIN DELLS
C O L U M B I A
PORTAGE
PARDEEVILLE
WYOCENA
FOX LAKE
BURNETT
BEAVER DAM
HORICON
JUNEAU
REEDSBURG
BARABOO
NORTH FREEDOM
S A U K
MERRIMAC
POYNETTE
RADIO STATION WIBU
FALL RIVER
COLUMBUS
CLYMAN
HUSTISFORD
R I C H L A N D
RICHLAND CENTER
LONE ROCK
PRAIRIE DU SAC AND SAUK CITY
LODI
REESEVILLE
D O D G E
DE FOREST
MUSCODA
AVOCA
SPRING GREEN
ARENA
MAZOMANIE
BLACK EARTH
WAUNAKEE
SUN PRAIRIE
WATERLOO
WATERTOWN
OCONOMOWOC
MIDDLETON
CROSS PLAINS
MADISON
COTTAGE GROVE
DEERFIELD
LAKE MILLS
JOHNSON'S CREEK
BARNEVELD
MT. HOREB
VERONA
MC FARLAND
RADIO STA. WIBU
CAMBRIDGE
JEFFERSON
SULLIVAN
I O W A
HIGHLAND
COBB
DODGEVILLE
RIDGEWAY
FENNIMORE
LINDEN
MINERAL POINT
DEWEY
OREGON
STOUGHTON
BELLEVILLE
BROOKLYN
BLANCHARDVILLE
EDGERTON
FT. ATKINSON
PALMYRA
J E F F E R S O N
D A N E

Fig. 1. Open-wire route from Madison to Stevens Point

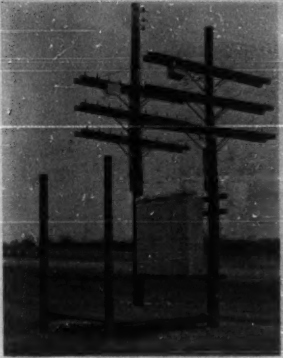


Fig. 2. Pole-mounted repeater installation

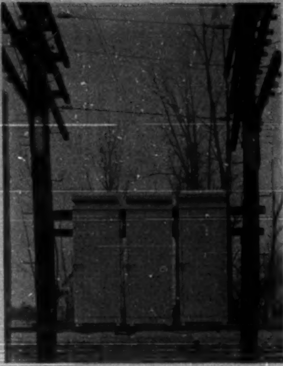


Fig. 3. Close-up view of repeater installation

normal condition can be determined from bridge measurements.

The repeaters for the *45A* system are not of the frequency-frogging type and therefore some means of preventing run-around crosstalk are necessary. A 10-foot gap is placed in the line as previously shown, and all non-*45A* pairs are brought into the cabinet, shown in Fig. 8, by means of 19-gauge paper-insulated cables. Protected terminals are used at each end of the cables. Each non-*45A* pair is equipped with a low-pass filter to provide attenuation in the range of frequencies amplified by the repeater. All wiring to the filters in the cabinet is shielded-pair wire. A 250-watt electric heating element controlled by a thermostat is provided to prevent moisture condensation in the cabinet.

Performance

Except for such factors as may be influenced by wide temperature variations, the performance of the regulating and amplifying circuits is the same for pole or building installation, and will not be discussed herein. The performance of the stand-by power supply and the crosstalk performance of the installation with its 10-foot gap in the line are of interest.

To determine the number of hours that the system would operate on the stand-by power unit, the transmitting and receiving sides of a channel were patched together at Stevens Point, and a transmission recording meter, set to record the round trip loss of the channel, was placed at the Madison terminal. Six and one-half hours after the a-c power was removed from a repeater, the system started to fail and was at complete failure within 10 minutes. This test was made at a battery temperature of 45 degrees and, during colder weather, a somewhat shorter time will be obtained. It is probably safe to consider that a 4-hour reserve is being obtained in this installation.

It should be pointed out that under power failure conditions the electric heater provided in the power cabinet is also without power. Hence there is danger of the batteries freezing during cold weather if they are allowed to go to complete discharge. Alarm circuits are provided to indicate when the repeater is operating on stand-by power. If conditions are such that a power failure may be of long duration, power may be restored at the repeater by the use of a small gasoline-driven generator such as provided to telephone construction forces for emergency lights and small power-operated equipment.

Transfers to the stand-by power supply and back again are made without noticeable reaction on either the transmission or the dial-signaling circuits on the system. No tests have been made with voice-frequency telegraph circuits to determine the effect on circuits of this type.

The crosstalk suppression obtained in the two installations in Wisconsin seems to be adequate for the operation of one system on the line. This system was installed with a minimum of line treatment. It is being operated on a phantomed transposition system designed for *C* car-

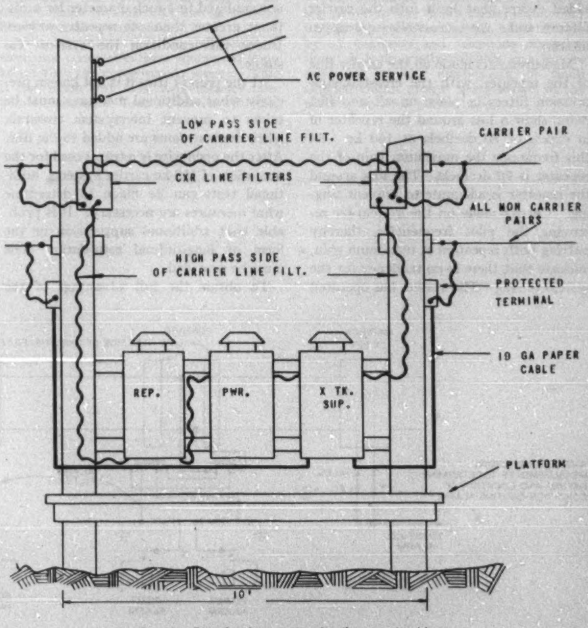


Fig. 4. Details of mounting, bridling, and cabling

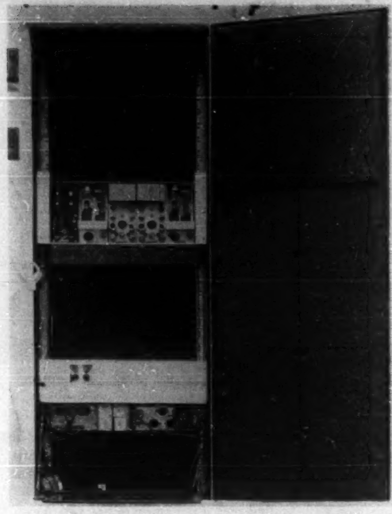


Fig. 5 (left). Interior of repeater cabinet showing complete repeater, top shelf, and power supply, lower shelf

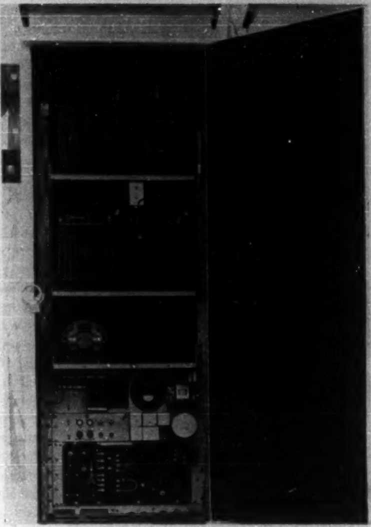


Fig. 6 (right). Stand-by power-supply unit, the same unit normally used in type-O carrier systems

rier. No longitudinal suppression is provided except that built into the carrier filters and the crosstalk-suppression filters.

Measurements made on the carrier line at the repeater, with the crosstalk-suppression filters in place on all non-*45A* pairs, show a loss around the repeater of in excess of 80 decibels at 150 kc. At this frequency the maximum gain of the repeater is 72 decibels. The loss around the repeater is adequate to prevent singing. Checks made on the system by removing the pilot frequencies, thereby putting both repeaters at maximum gain, indicate that there is no tendency for the system to sing. The system has operated under conditions during which hoarfrost accumulated to 1-inch diameter for a distance greater than one repeater section. Under this condition the system was stable.

At the present time it is not known precisely what additional measures must be taken to prevent intersystem crosstalk when more systems are added to the line. After the open wire is retransposed for the operation of 150-kc carrier systems, additional tests can be made to determine what measures are necessary. It is probable that additional suppression in the form of longitudinal retardation coils must be provided.

To obtain the full advantage of the crosstalk-suppression filters now installed, it seems certain that more attention must be paid to ground resistance than has been found necessary for the single system. The longitudinal suppression provided by the crosstalk-suppression filters depends to a considerable degree upon the resistance of the ground at the repeater installation. The ground as now provided consists of three 10-foot ground rods in sandy soil. No measurements of ground resistance have been made.

Tests made with the entire line sectionalized at the repeater point indicate that the 10-foot gap should be adequate. However, if it is not, it will be a relatively simple matter to extend one of the cables

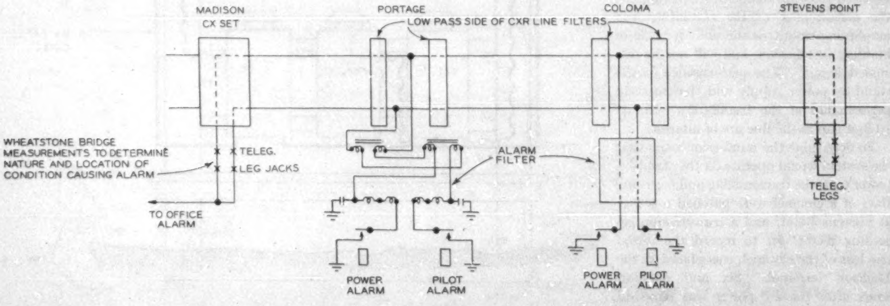

MADISON
CX SET
PORTAGE
LOW PASS SIDE OF CXR LINE FILTERS
COLOMA
STEVENS POINT
WHEATSTONE BRIDGE MEASUREMENTS TO DETERMINE NATURE AND LOCATION OF CONDITION CAUSING ALARM
TELEG. LEG JACKS
ALARM FILTER
TO OFFICE ALARM
TELEG. LEGS
POWER ALARM
PILOT ALARM
POWER ALARM
PILOT ALARM

Fig. 7 Schematic diagram of alarm circuit. Failure of a repeater will cause an alarm at Madison and Stevens Point

Fig. 8. Crosstalk suppression filter cabinet. All non-45A pairs are equipped with low-pass filters to provide attenuation in the range of frequencies amplified by the repeater

one span and to remove the wires in this span. If this is necessary, some arrangement of the carrier line filters may be required.

Conclusion

Pole mounting of 12-channel carrier systems seems to be practical and to offer some advantages in cost over a building when the number of systems is small. For a single system on the line, pole mounting of repeaters provides a means of obtaining additional circuits with a minimum of line treatment.

References

1. Type-O Carrier Telephone, J. A. Coy, E. K. Van Tassel. *AIEE Transactions*, vol. 71, pt. I, 1952 (Jan. 1953 section), pp. 428–37.

2. Application and Transmission Features of a New 12-Channel Open-Wire Carrier System, K. E. Appert, R. S. Caruthers, W. S. Cheskin. *AIEE Transactions*, vol. 73, pt. I, March 1954, pp. 18–27.

◆

Discussion

J. N. Petrie (Automatic Electric Company, Chicago, Ill): In this paper, the authors have done a good job of recording how they brushed aside a number of concepts which for many years had kept the application of high-frequency carrier systems on open-wire lines very much restricted. They are to be commended for having shaken off these theoretical shackles.

One of the purposes of this discussion is to point out that, while the emergency power supply arrangement referred to in the paper is absolutely essential for a pole-mounted repeater, it is nevertheless desirable to consider this approach as basic for all small repeater installations even if located in a telephone central office.

It is recognized that the reliability of electronic communications equipment is no better than the power supply source. This is the reason why most long-distance telephone equipment, such as carrier systems, has traditionally been operated from storage batteries. A battery plant to provide positive 130-volt plate battery together with its associated rectifiers and controls is an expensive item and could easily cost as much as the whole carrier repeater. From a practical point of view, an alternate solution is necessary in order to derive the economies of a-c operation and yet retain the reliability of a storage-battery power plant. This solution has been worked out as follows.

In a typical small telephone central office only a 48-volt battery is available. This battery serves at all times as the source of current for the vacuum-tube heaters while the 130-volt plate voltage would normally be derived directly from the 117-volt a-c mains by rectification.

In the event of the a-c main's failure, the 130-volt plate voltage would be derived from the 48-volt battery by means of a vibrator-type invertor and step-up transformer feeding the normal rectifier and filter network associated with the plate voltage supply.

Power units of this type are commercially available and have been applied in several repeater installations and in a considerable number of carrier terminal installations. The first cost of such a unit is about the same as a typical a-c power pack of the same capacity. The length of period for which emergency operation of a carrier repeater is practical with this equipment is limited only by the reserve capacity of the 48-volt storage battery.

This technique for economically supplying power to any small carrier equipment installation deserves serious consideration by all engineers and telephone organizations; it is especially valuable as a service safeguard in the establishing of small isolated circuit groups being leased to the military on a full-period basis.

LATE DISCUSSION

The following discussion was received too late to be included with the paper.

Minimizing and Sequential Circuits

Discussion of paper 55-436 by W. S. Bennett, published in Communications and Electronics, September 1955, pages 443–47.

John T. Blake and **John E. Schwenker** (Bell Telephone Laboratories, Inc., Murray Hill, N. J.): Mr. Bennett, by an unfortunate selection of examples, hides the inherent power of the Karnaugh-Veitch map. In the usual switching problem, the input signals are grounds appearing or disappearing on the input leads. To obtain a negation of an input therefore, one requires a primary relay; consequently, in normal terms, the example in Fig. 4 requires two relays (C and P) and Fig. 5 requires four (A, B, C, and Q) rather than one and three as stated. Sometimes it is possible to replace the primary relay with a multicontact switch, but this is an added bonus, and should not be assumed to be possible in a general method. The second Q contact in Fig. 5 presumably arises from this assumption, even though a 3-relay 7-square solution can be devised. (There is apparently a misprint in Fig. 5.)

Under the assumption of multicontact switches and $E=4$, equation 5 and the reasoning leading up to it are incorrect. The easiest way of showing this is by means of a counterexample (Fig. 10). In this example, E is 4, since there are four combinations of the two inputs; M is 8, since there are eight unique stable conditions. Equation 5 implies that $2^n-2n\geq 8-4$, or $n\geq 4$. The eight modes, however, map onto a 3-relay diagram, and a 3-relay circuit can be realized.

The corrected equation can be found as follows. It is necessary for the circuit to be able to distinguish between all of the unique stable states. It can distinguish between them because of different inputs (E), or because of different states of the internal memory relays (2^n), or a total of $E2^n$ distinguishable circuit states. Thus,

$$E2^n-2n\geq M \tag{35}$$

Even this equation does not determine n. If it is necessary for the circuit to operate from grounded leads, rather than switches, then perhaps as many as n relays will be required as determined from

$$2^n-2n\geq M \tag{36}$$

If the only difference between a state providing an output and another state is the presence of an input signal, then the state providing the output can often be eliminated, saving relays, as was done in Fig. 4. The n found from equation 36 may not be sufficient to make the mapping possible; for instance, n must be at least as large as the maximum number of ways of leaving a single state. The problem of determining the minimum number of relays to realize a given state diagram is topological rather than algebraic, and to our knowledge, no satisfactory method exists at present.

The discussion of the general problem of late-break contacts, or preliminary-make contacts is quite stimulating. Except for the special case of operating a relay through its own back contacts, the necessity of preliminary-make contacts is dependent on the contact network selected to realize the desired map. If a given mapping really requires a preliminary make, only remapping can eliminate it. The necessity for a preliminary-make can arise in the relay hold path each time an operated relay is supposed to hold while the circuit changes state; and in a shunt-down path when a relay is supposed to stay shunted down. Writing the canonical equations for overlapping subcubes, and providing the equivalent canonical contact network, will eliminate the necessity for preliminary-makes except for the afore-mentioned case. Often the network providing an overlapping subcube function will not require more contacts than one requiring preliminary-makes.* As an example, a more judicious arrangement of the contacts of Fig. 7 can eliminate the preliminary-make (Fig. 11). Shunt-down, as such, cannot eliminate any preliminary-make contacts.

Shunt-down is usually undesirable because of the necessity of providing the dropping resistor, because of the power loss in the dropping resistor, and because of the inherently slower operation. For these reasons, shunt-down cannot normally be justified by saving preliminary-makes, or even several contact pairs. Very often, however, shunt-down will save several relays, and can be justified on this basis. It does not seem necessary in the treatment of shunt-down, to depart from Boolean algebra by introducing special notation as in equation 32, or a special algebra such as Raspanti's. The success of these methods on Montgomerie's frequency-halving circuit is due to the obtaining of overlapping subcubes rather than the introduction of shunt-down. This can be shown by working out the correct Boolean expression for equation 33. The logical expression for $F_c(R)$ in Fig. 9 is: R is energized when either (r) or (P) are operated, and (a) and (P) are not both operated. In Boolean notation this becomes:

$$F_c(R)=(r+P)(aP)' \tag{37}$$

$$=(r+P)(a'+P') \tag{38}$$

$$=ra'+rP'+a'P+0 \tag{39}$$

Equation 39 clearly shows the overlapping subcubes; equation 38 was the one implemented in the improved circuit (Fig. 11). It would appear then that the need for, and the treatment of, preliminary-make contacts requires actual examination of the circuit or an algebra based on overlapping subcubes.

The minimization rule as stated by Mr. Bennett does not necessarily lead to the proper result, as we have shown. Transition time problems can be completely overcome by the use of overlapping subcubes except possibly where a relay must be operated through its own back contact. The map method itself provides a most practical means for utilizing shunt contact techniques. We feel that the map method is the most powerful and practical available tool for designing sequential circuits, and that certainly it is not limited to circuits of four or fewer relays.

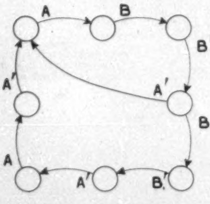


Fig. 10. Counterexample

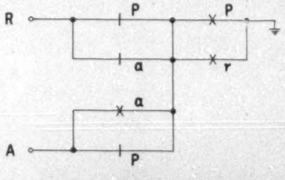


Fig. 11. Rearrangement of contacts of Fig. 7 of paper to eliminate preliminary-make

---

* It is possible to devise an algebra which preserves the overlapping subcubes, allowing circuit synthesis.

ERRATA

"Formulation of Magnetic Amplifier Circuit Equations" by D. H. Schaefer, published in Communication and Electronics, November 1955, pages 619–27.

1. Multiplication (X) signs at end of lines should be omitted in equations 18, 30, 34, 35, 64, 65, 104, 108, 110, 111, 112, 115, 119, 120, 123, 124, 125, and 133.

2. Equation 55 is of the following form:

$$\frac{A}{B}K_1-K_1I_1$$

where the numerator A is as printed, and the denominator B equals

$$(K_1+K_2+R_L)(K_1+K_2+R_c)-(K_1-K_2)^2$$

3. First symbol on third line of equation 112 should be a V symbol rather than a Λ symbol.

4. Equation 134 should read:

$$e=e_{ac}+Ne_cV[Ne_c\Lambda\alpha e_{ac}+\beta Ne_c\Lambda e_{ac}+\gamma Ne_c]$$

as

$$\alpha<1,\ \gamma<1,\ \mu<1,\text{ and }\phi>\gamma>\beta$$